

# Contents

<b>9</b>	<b>Interference and Coherence</b>	<b>1</b>
9.1	Overview . . . . .	1
9.2	Coherence . . . . .	2
9.2.1	Young’s Slits . . . . .	2
9.2.2	Interference with an Extended Source: van Cittert-Zernike Theorem . . . . .	4
9.2.3	More General Formulation of Spatial Coherence; Lateral Coherence Length . . . . .	7
9.2.4	Generalization to Two Dimensions . . . . .	8
9.2.5	Michelson Stellar Interferometer, and <b>T2</b> Box on Atmospheric Seeing . . . . .	9
9.2.6	Temporal Coherence . . . . .	15
9.2.7	Michelson Interferometer and Fourier-Transform Spectroscopy . . . . .	17
9.2.8	Degree of Coherence; Relation to Theory of Random Processes . . . . .	20
9.3	Radio Telescopes . . . . .	22
9.3.1	Two-Element Radio Interferometer . . . . .	23
9.3.2	Multiple Element Radio Interferometer . . . . .	23
9.3.3	Closure Phase . . . . .	24
9.3.4	Angular Resolution . . . . .	25
9.4	Etalons and Fabry-Perot Interferometers . . . . .	26
9.4.1	Multiple Beam Interferometry; Etalons . . . . .	26
9.4.2	Fabry-Perot Interferometer, and Box on Modes of a Fabry-Perot Cavity with Spherical Mirrors . . . . .	33
9.4.3	Fabry-Perot Applications: spectrometer, laser, mode-cleaning cavity, beam-shaping cavity, PDH laser stabilization, optical frequency comb . . . . .	37
9.5	<b>T2</b> Laser Interferometer Gravitational Wave Detectors . . . . .	43
9.6	Power Correlations and Photon Statistics: Hanbury Brown & Twiss Intensity Interferometer . . . . .	49

# Chapter 9

## Interference and Coherence

Version 1209.1.K.pdf, 12 December 2012

*Please send comments, suggestions, and errata via email to [kip@caltech.edu](mailto:kip@caltech.edu) or on paper to Kip Thorne, 350-17 Caltech, Pasadena CA 91125*

### Box 9.1 Reader's Guide

- This chapter depends substantially on
  - Secs. 8.2, 8.3 and 8.5.5 of Chap. 8
  - Correlation functions, spectral densities and the Wiener-Khintchine theorem for random processes, Sec. 6.4 of Chap. 6.
- The concept of coherence length or coherence time, as developed in this chapter, will be used in Chaps. 10, 15, 16 and 23 of this book.
- Interferometry as developed in this chapter, especially in Sec. 9.5, is a foundation for the discussion of gravitational-wave detection in Sec. 27.6.
- Nothing else in this book relies substantially on this chapter.

### 9.1 Overview

In the last chapter, we considered superpositions of waves that pass through a (typically large) aperture. The foundation for our analysis was the Helmholtz-Kirchoff expression (8.4) for the field at a chosen point  $\mathcal{P}$  as a sum of contributions from all points on a closed surface surrounding  $\mathcal{P}$ . The spatially varying field pattern resulting from this superposition of many different contributions was called *diffraction*.

In this chapter, we continue our study of superposition, but for the more special case where only two or at most several discrete beams are being superposed. For this special case one uses the term *interference* rather than diffraction. Interference is important in a wide variety of practical instruments designed to measure or utilize the spatial and temporal structures of electromagnetic radiation. However interference is not just of practical importance. Attempting to understand it forces us to devise ways of describing the radiation field that are independent of the field's origin and independent of the means by which it is probed. Such descriptions lead us naturally to the fundamental concept of *coherence* (Sec. 9.2).

The light from a distant, monochromatic point source is effectively a plane wave; we call it “perfectly coherent” radiation. In fact, there are two different types of coherence present: *lateral or spatial coherence* (coherence in the angular structure of the radiation field), and *temporal or longitudinal coherence* (coherence in the field's temporal structure, which clearly must imply something also about its frequency structure). We shall see in Sec. 9.2 that for both types of coherence there is a measurable quantity, called the *degree of coherence*, that is the Fourier transform of either the radiation's angular intensity distribution  $I(\boldsymbol{\alpha})$  (energy flux per unit angle or solid angle, as a function of direction  $\boldsymbol{\alpha}$ ) or its spectral energy flux  $F_\omega(\omega)$  (energy flux per unit angular frequency  $\omega$ , as a function of angular frequency).

Interspersed with our development of the theory of coherence are two applications: (i) the *stellar interferometer* (Sec. 9.2.5), by which Michelson measured the diameters of Jupiter's moons and several bright stars using spatial coherence; and (ii) the *Michelson interferometer* and its practical implementation in a *Fourier-transform spectrometer* (Sec. 9.2.7), which use temporal coherence to measure electromagnetic spectra, e.g. the spectral energy flux of the cosmic microwave background radiation. After developing our full formalism for coherence, we shall go on in Sec. 9.3 to apply it to the operation of *radio telescopes*, which function by measuring the spatial coherence of the radiation field.

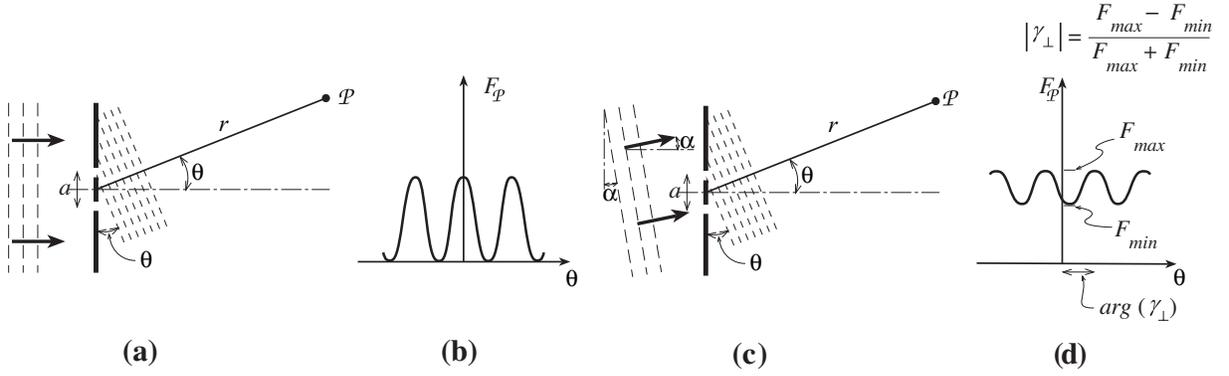
In Sec. 9.4 we shall turn to multiple beam interferometry, in which incident radiation is split many times into several different paths and then recombined. A simple example is an *etalon*, made from two parallel, reflecting surfaces. A Fabry-Perot cavity interferometer, in which light is trapped between two highly reflecting mirrors (e.g. in a laser), is essentially a large scale etalon. In Secs. 9.4.3 and 9.5 we discuss a number of applications of Fabry Perot interferometers, including *lasers, their stabilization, manipulations of laser light, the optical frequency comb, and laser interferometer gravitational wave detectors*.

Finally, in Sec. 9.6, we shall turn to the *intensity interferometer*, which although it has not proved especially powerful in application, does illustrate some quite subtle issues of physics and, in particular, highlights the relationship between the classical and quantum theories of light.

## 9.2 Coherence

### 9.2.1 Young's Slits

The most elementary example of interference is provided by Young's slits. Suppose two long, narrow, parallel slits are illuminated coherently by *monochromatic light* from a distant source that lies on the perpendicular bisector of the line joining the slits (the *optic axis*),



**Fig. 9.1:** (a) Young's Slits. (b) Interference fringes observed in a transverse plane [Eq. (9.1b)]. (c) The propagation direction of the incoming waves is rotated to make an angle  $\alpha$  to the optic axis; as a result, the angular positions of the interference fringes in drawing (b) are shifted by  $\Delta\theta = \alpha$  [Eq. (9.3); not shown]. (d) Interference fringes observed from an extended source [Eq. (9.8)].

so an incident wavefront reaches the slits simultaneously (Fig. 9.1a). This situation can be regarded as having only one lateral dimension because of translation invariance in the other. The waves from the slits (effectively, two one-dimensional beams) fall onto a screen in the distant, Fraunhofer region, and there they interfere. The Fraunhofer interference pattern observed at a point  $\mathcal{P}$ , whose position is specified using the polar coordinates  $(r, \theta)$  shown in Fig. 9.1, is proportional to the spatial Fourier transform of the transmission function [Eq. (8.11a)]. If the slits are very narrow, we can regard the transmission function as two  $\delta$ -functions, separated by the slit spacing  $a$ , and its Fourier transform will be

$$\psi(\theta) \propto e^{-ika\theta/2} + e^{ika\theta/2} \propto \cos\left(\frac{ka\theta}{2}\right), \quad (9.1a)$$

where  $k = 2\pi/\lambda$  is the light's wave number and  $a$  is the slit's separation. (That we can sum the wave fields from the two slits in this manner is a direct consequence of the linearity of the underlying wave equation.) The *energy flux* (energy per unit time crossing a unit area) at  $\mathcal{P}$  (at angle  $\theta$  to the optic axis) will be

$$\boxed{F(\theta) \propto |\psi|^2 \propto \cos^2(ka\theta/2)}; \quad (9.1b)$$

cf. Fig. 9.1b. The alternating regions of dark and bright illumination in this flux distribution are known as *interference fringes*. Notice that the flux falls to zero between the bright fringes. This will be very nearly so even if (as is always the case in practice) the field is very slightly non-monochromatic, i.e. even if the field hitting the slits has the form  $e^{i[\omega_0 t + \delta\varphi(t)]}$ , where  $\omega_0 = c/k$  is the light's average angular frequency and  $\delta\varphi(t)$  is a phase [not to be confused with the light's full phase  $\varphi = \omega t + \delta\varphi(t)$ ], which varies randomly on a timescale extremely long compared to  $1/\omega_0$ .<sup>1</sup> Notice also that there are many fringes, symmetrically disposed with respect to the optic axis. [If we were to take account of the finite width  $w \ll a$  of

<sup>1</sup>More precisely, if  $\delta\varphi(t)$  wanders by  $\sim \pi$  on a timescale  $\tau_c \gg 2\pi/\omega_0$  (the waves' *coherence time*), then the waves are contained in a bandwidth  $\Delta\omega_0 \sim 2\pi/\tau_c \ll \omega_0$  centered on  $\omega_0$ ,  $k$  is in a band  $\Delta k \sim k\Delta\omega/\omega_0$ ,

the two slits, then we would find, by contrast with Eq. (9.1b) that the actual number of fringes is finite, in fact of order  $a/w$ ; cf. Fig. 8.6 and associated discussion.] This type of interferometry is sometimes known as *interference by division of the wave front*.

This Young's slits experiment is, of course, familiar from quantum mechanics, where it is often used as a striking example of the non-particulate behavior of electrons.<sup>2</sup> Just as for electrons, so also for photons, it is possible to produce interference fringes even if only one photon is in the apparatus at any time, as was demonstrated in a famous experiment performed by G. I. Taylor in 1909. However, our concerns in this chapter are with the classical limit, where many photons are present simultaneously and their fields can be described by Maxwell's equations. In the next subsection we shall depart from the usual quantum mechanical treatment by asking what happens to the fringes when the source of radiation is spatially extended.

\*\*\*\*\*

## EXERCISES

### Exercise 9.1 *Problem: Single Mirror Interference*

X-rays with wavelength  $8.33\text{\AA}$  ( $0.833\text{ nm}$ ) coming from a point source can be reflected at shallow angles of incidence from a plane mirror. The direct ray from a point source to a detector  $3\text{ m}$  away interferes with the reflected ray to produce fringes with spacing  $25\mu\text{m}$ . Calculate the distance of the X-ray source from the mirror plane.

\*\*\*\*\*

## 9.2.2 Interference with an Extended Source: van Cittert-Zernike Theorem

We shall approach the topic of extended sources in steps. Our first step was taken in the last subsection, where we dealt with an idealized, single, incident plane wave, such as might be produced by an ideal, distant laser. We have called this type of radiation *perfectly coherent*, which we have implicitly taken to mean that the field oscillates with a fixed angular frequency  $\omega_o$  and a randomly but very slowly varying phase  $\delta\varphi(t)$  (see footnote 1), and thus, for all practical purposes, there is a time-independent phase difference between any two points within the region under consideration.

As our second step, we keep the incoming waves perfectly coherent and perfectly planar, but change their incoming direction in Fig. 9.1 so it makes a small angle  $\alpha$  to the optic axis (and correspondingly its wave fronts make an angle  $\alpha$  to the plane of the slits) as shown in

---

and the resulting superposition of precisely monochromatic waves has fringe minima with fluxes  $F_{\min}$  that are smaller than the maxima by  $F_{\min}/F_{\max} \sim (\pi\Delta\omega/\omega_o)^2 \ll 1$ . (One can see this in order of magnitude by superposing the flux (9.1b) with wave number  $k$  and the same flux with wave number  $k + \Delta k$ .) Throughout this section, until Eq. (9.15) we presume that the waves have such a small bandwidth (such a long coherence time) that this  $F_{\min}/F_{\max}$  is completely negligible; for example,  $1 - F_{\min}/F_{\max}$  is far closer to unity than any fringe visibility  $V$  [Eq. (9.8) below] that is of interest to us. This can be achieved in practice by either controlling the waves' source, or by band-pass filtering the measured signals just before detecting them.

<sup>2</sup>See, e.g., Chapter 1 of Volume III of Feynman, Leighton, and Sands (1965).

Fig. 9.1c. Then the distribution of energy flux in the Fraunhofer diffraction pattern on the screen will be modified to

$$\begin{aligned} F(\theta) &\propto |e^{-ika(\theta-\alpha)/2} + e^{+ika(\theta-\alpha)/2}|^2 \propto \cos^2\left(\frac{ka(\theta-\alpha)}{2}\right) \\ &\propto \{1 + \cos[ka(\theta-\alpha)]\}. \end{aligned} \quad (9.2)$$

Notice that, as the direction  $\alpha$  of the incoming waves is varied, the locations of the bright and dark fringes change by  $\Delta\theta = \alpha$ , but the fringes remain fully sharp (their minima remain essentially zero; cf. footnote 1). Thus, *the positions of the fringes carry information about the direction to the source.*

Now, in our third and final step, we will deal with an extended source, i.e. one whose radiation comes from a finite range of angles  $\alpha$ , with (for simplicity)  $|\alpha| \ll 1$ . We shall assume that the source is monochromatic (and in practice we can make it very nearly monochromatic by band-pass filtering the waves just before detection). However, in keeping with how all realistic monochromatic sources (including band-pass filtered sources) behave, we shall give it a randomly fluctuating phase  $\delta\varphi(t)$  [and amplitude  $A(t)$ ], and shall require that the timescale on which the phase and amplitude wander (the waves' coherence time) be very long compared to the waves' period  $2\pi/\omega_o$ ; cf. footnote 1.

We shall also *assume that, as for almost all realistic sources, the fluctuating phases in the waves from different directions are completely uncorrelated.* To make this precise, we write the field in the form<sup>3</sup>

$$\Psi(x, z, t) = e^{i(kz - \omega_o t)} \int \psi(\alpha, t) e^{ik\alpha x} d\alpha, \quad (9.3)$$

where  $\psi(\alpha, t) = Ae^{-i\delta\varphi}$  is the slowly wandering complex amplitude of the waves from direction  $\alpha$ . When we consider the total flux arriving at a given point  $(x, z)$  from two different directions  $\alpha_1$  and  $\alpha_2$  and average it over times long compared to the waves' coherence time, then we lose all interference between the two contributions:

$$\overline{|\psi(\alpha_1, t) + \psi(\alpha_2, t)|^2} = \overline{|\psi(\alpha_1, t)|^2} + \overline{|\psi(\alpha_2, t)|^2}. \quad (9.4)$$

Such radiation is said to be *incoherent in the incoming angle*  $\alpha$ , and we say that the contributions from different directions superpose incoherently. This is just a fancy way of saying that their intensities (averaged over time) add linearly.

The angularly incoherent light from our extended source is sent through two Young's slits and produces fringes on a screen in the distant Fraunhofer region. We assume that the coherence time for the light from each source point is very long compared to the difference in light travel time to the screen via the two different slits. Then the light from each source point in the extended source forms the sharp interference fringes described by Eq. (9.2). However, because contributions from different source directions add incoherently, the flux distribution on the screen is a linear sum of the fluxes from all the source points:

$$F(\theta) \propto \int d\alpha I(\alpha) \{1 + \cos[ka(\theta - \alpha)]\} \quad (9.5)$$

---

<sup>3</sup>As in Chap. 8, we denote the full field by  $\Psi$  and reserve  $\psi$  to denote the portion of  $\Psi$  the field from which a monochromatic part  $e^{-i\omega_o t}$  or  $e^{i(kz - \omega_o t)}$  has been factored out.

Here  $I(\alpha)d\alpha \propto |\overline{\psi(\alpha, t)}|^2 d\alpha$  is the flux incident on the plane of the slits from the infinitesimal range  $d\alpha$  of directions, i.e.  $I(\alpha)$  is the radiation's *intensity* (its energy per unit time falling onto a unit area and coming from a unit angle). The remainder of the integrand,  $1 + \cos[ka(\theta - \alpha)]$ , is the Fraunhofer diffraction pattern (9.2) for coherent radiation from direction  $\alpha$ .

We presume that *the range of angles present in the waves,  $\Delta\alpha$ , is large compared to their fractional bandwidth  $\Delta\alpha \gg \Delta\omega/\omega_o$* ; so, whereas the finite but tiny bandwidth produced negligible smearing out of the interference fringes (footnote 1), the finite but small range of directions may produce significant smearing, i.e. the minima of  $F(\theta)$  might not be very sharp. We quantify the fringes' non-sharpness and their locations by writing the slit-produced flux distribution (9.5) in the form

$$F(\theta) = F_S [1 + \Re\{\gamma_{\perp}(ka)e^{-ika\theta}\}], \quad (9.6a)$$

where

$$F_S \equiv \int d\alpha I(\alpha) \quad (9.6b)$$

(subscript  $S$  for "source") is the total flux arriving at the slits from the source, and

$$\gamma_{\perp}(ka) \equiv \frac{\int d\alpha I(\alpha)e^{ika\alpha}}{F_S} \quad (9.7a)$$

is known as the radiation's *degree of spatial (or lateral) coherence*. The phase of  $\gamma_{\perp}$  determines the angular locations of the fringes; its modulus determines their depth (the amount of their smearing due to the source's finite angular size).

*The nonzero value of  $\gamma_{\perp}(ka)$  reflects the fact that there is some amount of relative coherence between the waves arriving at the two slits, whose separation is  $a$ .* The radiation can have this finite spatial coherence, despite its complete lack of angular coherence, because each angle contributes coherently to the field at the two slits. The lack of coherence for different angles reduces the net spatial coherence (smears the fringes), but does not drive the coherence all the way to zero (does not completely destroy the fringes).

Eq. (9.7a) says that *the degree of spatial coherence of the radiation from an extended, angularly incoherent source is the Fourier transform of the source's angular intensity pattern*. Correspondingly, if one knows the degree of spatial coherence as a function of the (dimensionless) distance  $ka$ , from it one can reconstruct the source's angular intensity pattern by Fourier inversion:

$$I(\alpha) = F_S \int \frac{d(ka)}{2\pi} \gamma_{\perp}(ka)e^{-ika\alpha}. \quad (9.7b)$$

The two Fourier relations (9.7a), (9.7b) are called the *van Cittert-Zernike Theorem*. In Ex. 9.8, we shall see that this theorem is a complex-variable version of Chap. 6's *Wiener-Khintchine Theorem* for random processes.

Because of its Fourier-transform relationship to the source's angular intensity pattern  $I(\alpha)$ , the degree of spatial coherence  $\gamma_{\perp}(ka)$  is of great practical importance. *For a given*

choice of  $ka$  (a given distance between the slits),  $\gamma_{\perp}$  is a complex number that one can read off the interference fringes of Eq. (9.6a) and Fig. 9.1d as follows: Its modulus is

$$\boxed{|\gamma_{\perp}| \equiv V = \frac{F_{\max} - F_{\min}}{F_{\max} + F_{\min}}} \quad (9.8)$$

where  $F_{\max}$  and  $F_{\min}$  are the maximum and minimum values of the flux  $F$  on the screen; and its phase  $\arg(\gamma_{\perp})$  is  $ka$  times the displacement  $\Delta\theta$  of the centers of the bright fringes from the optic axis. The modulus is called the *fringe visibility*, or simply the *visibility*, because of its measuring the fractional contrast in the fringes [Eq. (9.8)], and this name is the reason for the symbol  $V$ . Analogously, the complex quantity  $\gamma_{\perp}$  (or a close relative) is sometimes known as the *complex fringe visibility*. Notice that  $V$  can lie anywhere in the range from zero (no contrast; fringes completely undetectable) to unity (monochromatic plane wave; contrast as large as possible). When the phase  $\arg(\gamma_{\perp})$  of the complex visibility (degree of coherence) is zero, there is a bright fringe precisely on the optic axis. This will be the case, e.g., for a source that is symmetric about the optic axis. If the symmetry point of such a source is gradually moved off the optic axis by an angle  $\delta\alpha$ , the fringe pattern will shift correspondingly by  $\delta\theta = \delta\alpha$ , and this will show up as a corresponding shift in the argument of the fringe visibility,  $\arg(\gamma_{\perp}) = ka\delta\alpha$ .

The above analysis shows that Young's slits are nicely suited to measuring both the modulus and the phase of the complex fringe visibility (the degree of spatial coherence) of the radiation from an extended source.

### 9.2.3 More General Formulation of Spatial Coherence; Lateral Coherence Length

It is not necessary to project the light onto a screen to determine the contrast and angular positions of the fringes. For example, if we had measured the field at the locations of the two slits, we could have combined the signals electronically and cross correlated them numerically to determine what the fringe pattern would be with slits. All we are doing with the Young's slits is sampling the wave field at two different points, which we now shall label 1 and 2. Observing the fringes corresponds to adding a phase  $\varphi$  ( $= ka\theta$ ) to the field at one of the points and then adding the fields and measuring the flux  $\propto |\psi_1 + \psi_2 e^{i\varphi}|^2$  averaged over many periods. Now, since the source is far away, the rms value of the wave field will be the same at the two slits,  $\overline{|\psi_1|^2} = \overline{|\psi_2|^2} \equiv \overline{|\psi|^2}$ . We can therefore express this time averaged flux in the symmetric-looking form

$$\begin{aligned} F(\varphi) &\propto \overline{(\psi_1 + \psi_2 e^{i\varphi})(\psi_1^* + \psi_2^* e^{-i\varphi})} \\ &\propto 1 + \Re \left( \frac{\overline{\psi_1 \psi_2^*}}{\overline{|\psi|^2}} e^{-i\varphi} \right). \end{aligned} \quad (9.9)$$

Here a bar denotes an average over times long compared to the coherence times for  $\psi_1$  and  $\psi_2$ . Comparing with Eq. (9.6a) and using  $\varphi = ka\theta$ , we identify

$$\boxed{\gamma_{\perp 12} = \frac{\overline{\psi_1 \psi_2^*}}{\overline{|\psi|^2}}} \quad (9.10)$$

as the *degree of spatial coherence* in the radiation field between the two points 1 and 2. Equation (9.10) is the general definition of degree of spatial coherence. Equation (9.6a) is the special case for points separated by a lateral distance  $a$ .

If the radiation field is strongly correlated between the two points, we describe it as having strong *spatial* or *lateral coherence*. Correspondingly, we shall define a field's *lateral coherence length*  $l_{\perp}$  as the linear size of a region over which the field is strongly correlated (has  $V = |\gamma_{\perp}| \sim 1$ ). *If the angle subtended by the source is  $\sim \delta\alpha$ , then by virtue of the van Cittert-Zernike theorem (9.7) and the usual reciprocal relation for Fourier transforms, the radiation field's lateral coherence length will be*

$$\boxed{l_{\perp} \sim \frac{2\pi}{k \delta\alpha} = \frac{\lambda}{\delta\alpha}}. \quad (9.11)$$

This relation has a simple physical interpretation. Consider two beams of radiation coming from opposite sides of the brightest portion of the source. These beams will be separated by the incoming angle  $\delta\alpha$ . As one moves laterally in the plane of the Young's slits, one will see a varying relative phase delay between these two beams. The coherence length  $l_{\perp}$  is the distance over which the variations in that relative phase delay are of order  $2\pi$ ,  $k \delta\alpha l_{\perp} \sim 2\pi$ .

## 9.2.4 Generalization to Two Dimensions

We have so far just considered a one-dimensional intensity distribution  $I(\alpha)$  observed through the familiar Young's slits. However, most sources will be two dimensional, so in order to investigate the full radiation pattern, we should allow the waves to come from 2-dimensional angular directions  $\boldsymbol{\alpha}$ , whence

$$\Psi = e^{i(kz - \omega_0 t)} \int \psi(\boldsymbol{\alpha}, t) e^{ik\boldsymbol{\alpha} \cdot \mathbf{x}} d^2\boldsymbol{\alpha} \equiv e^{i(kz - \omega_0 t)} \psi(\mathbf{x}, t) \quad (9.12a)$$

[where  $\psi(\boldsymbol{\alpha}, t)$  is slowly varying]; and we should use several pairs of slits aligned along different directions. Stated more generally, we should sample the wave field (9.12a) at a variety of points separated by a variety of two-dimensional vectors  $\mathbf{a}$  transverse to the direction of wave propagation. *The complex visibility (degree of spatial coherence)* will then be a function of  $k\mathbf{a}$ ,

$$\boxed{\gamma_{\perp}(k\mathbf{a}) = \frac{\overline{\psi(\mathbf{x}, t)\psi^*(\mathbf{x} + \mathbf{a}, t)}}{|\overline{\psi}|^2}}, \quad (9.12b)$$

and the van Cittert-Zernike Theorem (9.7) (actually the Wiener-Khintchine theorem in disguise; see Ex. 9.8) will take the two-dimensional form

$$\boxed{\gamma_{\perp}(k\mathbf{a}) = \frac{\int d\Omega_{\alpha} I(\boldsymbol{\alpha}) e^{ik\mathbf{a} \cdot \boldsymbol{\alpha}}}{F_S}}, \quad (9.13a)$$

$$\boxed{I(\boldsymbol{\alpha}) = F_S \int \frac{d^2(k\mathbf{a})}{(2\pi)^2} \gamma_{\perp}(k\mathbf{a}) e^{-ik\mathbf{a} \cdot \boldsymbol{\alpha}}}. \quad (9.13b)$$

Here  $I(\boldsymbol{\alpha}) \propto |\overline{\psi(\boldsymbol{\alpha}, t)}|^2$  is the source's *intensity* (energy per unit time crossing a unit area from a unit solid angle  $d\Omega_\alpha$ ;  $F_S = \int d\Omega_\alpha I(\boldsymbol{\alpha})$  is the source's total energy flux; and  $d^2(ka) = k^2 d\Sigma_a$  is a (dimensionless) surface area element in the lateral plane.

\*\*\*\*\*

## EXERCISES

**Exercise 9.2** *Problem: Lateral Coherence of solar radiation*

How closely separated must a pair of Young's slits be to see strong fringes from the sun (angular diameter  $\sim 0.5^\circ$ ) at visual wavelengths? Suppose that this condition is just satisfied and the slits are  $10\mu\text{m}$  in width. Roughly how many fringes would you expect to see?

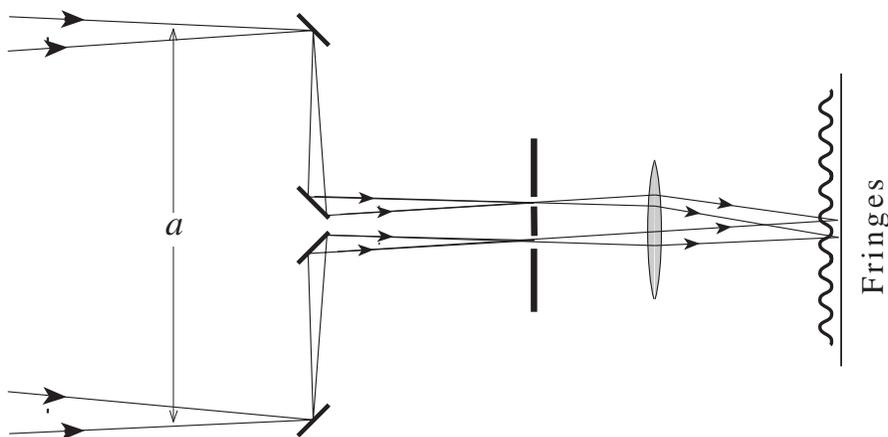
**Exercise 9.3** *Problem: Degree of Coherence for a Source with Gaussian Intensity Distribution*

A circularly symmetric light source has an intensity distribution  $I(\boldsymbol{\alpha}) = I_0 \exp(-\alpha^2/2\alpha_0^2)$ , where  $\alpha$  is the angular radius measured from the optic axis. Compute the degree of spatial coherence. What is the lateral coherence length? What happens to the degree of spatial coherence and the interference fringe pattern if the source is displaced from the optic axis?

\*\*\*\*\*

### 9.2.5 Michelson Stellar Interferometer, and T2 Box on Atmospheric Seeing

The classic implementation of Young's slits for measuring spatial coherence is Michelson's stellar interferometer, which Albert A. Michelson and Francis Pease used for measuring the angular diameters Betelgeuse and several other bright stars in 1920. The star light was sampled at two small mirrors separated by a variable distance  $a \leq 6\text{m}$  and was then reflected



**Fig. 9.2:** Schematic Illustration of a Michelson Stellar Interferometer.

into the 100 inch (2.5 meter) telescope on Mount Wilson, California, to form interference fringes; cf. Fig. 9.2. (As we have emphasized, the way in which the fringes are formed is unimportant; all that matters is the two locations where the light is sampled, i.e. the first two mirrors in Fig. 9.2.) As Michelson and Pease increased the separation  $a$  between the mirrors, the fringe visibility  $V$  decreased. Michelson and Pease modeled Betelgeuse (rather badly in fact) as a circular disk of uniform brightness,  $I(\boldsymbol{\alpha}) = 1$  for  $|\boldsymbol{\alpha}| < \alpha_r$  and 0 for  $|\boldsymbol{\alpha}| > \alpha_r$ , so its visibility was given, according to Eq. (9.13a), as

$$V = \gamma_{\perp} = 2\text{jinc}(ka\alpha_r) \quad (9.14)$$

where  $\alpha_r$  is the star's true angular radius and  $\text{jinc}(\xi) = J_1(\xi)/\xi$ . They identified the separation  $a \simeq 3$  m, where the fringes disappeared, with the first zero of the function  $\text{jinc}(ka)$ , and they thereby inferred that the angular radius of Betelgeuse is  $\alpha_r \sim 0.02$  arc seconds, which at Betelgeuse's (parallax-measured) distance of 200pc (600 lyr) corresponds to a physical radius  $\sim 300$  times larger than that of the Sun, a reasonable value in light of the modern theory of stellar structure.

This technique only works for big, bright stars and is very difficult to use because turbulence in the atmosphere causes the fringes to keep moving about; see Box 9.2 and Ex. 9.4 for details.

\*\*\*\*\*

## EXERCISES

**Exercise 9.4** T2 *Example and Derivation: Time-Averaged Visibility and Image for a Distant Star Seen Through Earth's Turbulent Atmosphere*

Fill in the details of the analysis of time-averaged seeing in Box 9.2. More specifically:

- (a) Give an order-of-magnitude derivation of Eq. (4a) for the mean-square phase fluctuation of light induced by propagation through a thin, turbulent layer of atmosphere. [Hint: consider turbulent cells of size  $a$ , each of which produces some  $\delta\varphi$ , and argue that the contributions add up as a random walk.]
- (b) Deduce the factor 2.91 in Eq. (4a) by evaluating  $D_{\delta\varphi} = k^2 \left\langle \left\{ \int_z^{z+\delta h} [\delta\mathbf{n}(\mathbf{x} + a, z, t) - \delta\mathbf{n}(\mathbf{x}, z, t)] dz \right\}^2 \right\rangle$ .
- (c) Derive Eq. (4b) for the time-averaged complex visibility after propagating through the thin layer. [Hint: argue that, because  $\zeta \equiv \delta\varphi(\mathbf{x}, t) - \delta\varphi(\mathbf{x} + \mathbf{a}, t)$  is the result of contributions from a huge number of independent turbulent cells, the central limit theorem (Sec. 6.3.2) guarantees it is a Gaussian random variable. Then evaluate  $\gamma_{\perp} = \langle e^{i\zeta} \rangle = \int_{-\infty}^{\infty} p(\zeta) e^{i\zeta} d\zeta$  with  $p(\zeta)$  the Gaussian distribution.]
- (d) Use the point-spread function (8.28) for free propagation of the light field  $\psi$  to show that, under free propagation, the complex visibility  $\gamma_{\perp}(k\mathbf{a}, z, t) = \langle \psi(\mathbf{x} + \mathbf{z}, z, t) \psi^*(\mathbf{x}, z, t) \rangle$  (with averaging over  $\mathbf{x}$  and  $t$ ) is constant, i.e. independent of height  $z$ .

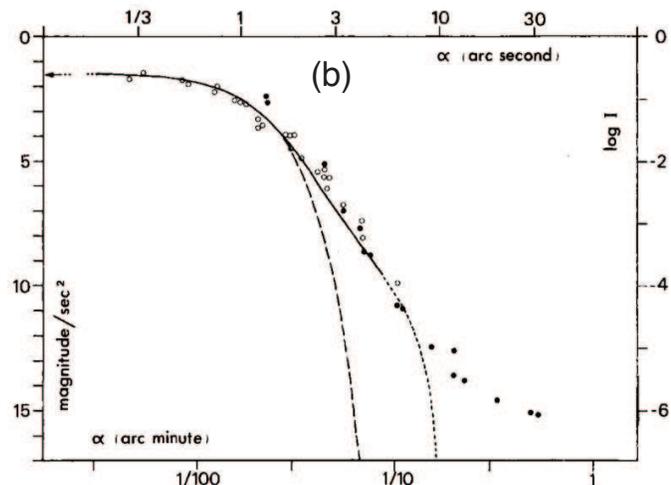
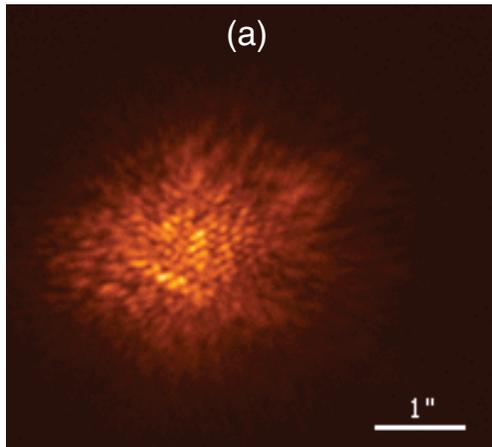
## Box 9.2

**T2** Astronomical Seeing, Speckel Image Processing, and Adaptive Optics

When light from a star passes through turbulent layers of the earth's atmosphere, the turbulently varying index of refraction  $n(\mathbf{x}, t)$  diffracts the light in a random, time varying way. One result is "twinkling" (fluctuations in the flux observed by eye on the ground, with fluctuational frequencies  $\sim 100$  Hz). Another is astronomical seeing: the production of many images of the star, i.e. *speckles*, as seen through a large optical telescope [Fig. (a) below], with the image pattern fluctuating at  $\sim 100$  Hz.

In this box and Ex. 9.4, we shall quantify astronomical seeing using the theory of two-dimensional lateral coherence. We do this not because seeing is important (though it is), but rather because our analysis provides an excellent illustration of three fundamental concepts working together: (i) turbulence in fluids, and its Kolmogorov spectrum (Chap. 15); (ii) random processes (Chap. 6), and (iii) coherence of light (this chapter).

We shall begin by deriving, for a star with arbitrarily small angular diameter, the time-averaged complex visibility  $\gamma_{\perp}$  observed on the ground, and the visibility's Fourier transform, the observed intensity distribution averaged over the speckles,  $\bar{I}(\boldsymbol{\alpha})$ . Then we shall briefly discuss the temporally fluctuating speckel pattern and techniques for extracting information from it.



Figures: (a) Picture of a bright star with a dimmer companion star, as seen through the Russian 6 meter telescope in an exposure shorter than 10 ms. Atmospheric turbulence creates a large number of images of each star (speckles) spread over a region with angular diameter of order 2 arc seconds. (b) The theory discussed in the text and in Ex. 9.4 predicts the solid curve for the time averaged intensity distribution, in the case of a single bright star. Notice the logarithmic axes. The dotted curve is an estimate of the influence of the small-scale cutoff of the turbulence, and the dashed curve is a Gaussian. The circles are observational data. [(a) is from the infrared astronomy group of the Max Planck Institut for Radioastronomy; (b) is from Roddier (1981).]

**Box 9.2 (Continued)**

***Time-Averaged Visibility and Angular Distribution of Intensity***

When analyzing light propagation through the turbulent atmosphere, it is convenient to describe the turbulent fluctuations of the index of refraction by their spatial correlation function  $C_n(\boldsymbol{\xi}) \equiv \langle \delta \mathbf{n}(\mathbf{X}, t) \delta \mathbf{n}(\mathbf{X} + \boldsymbol{\xi}, t) \rangle$  (discussed in Sec. 6.4.1); or, better yet, by  $\mathbf{n}$ 's *mean-square fluctuation on the lengthscale*  $\xi$ ,

$$D_n(\boldsymbol{\xi}) \equiv \langle [\delta \mathbf{n}(\mathbf{X} + \boldsymbol{\xi}, t) - \delta \mathbf{n}(\mathbf{X}, t)]^2 \rangle = 2[\sigma_n^2 - C_n(\boldsymbol{\xi})], \quad (1)$$

which is called  $\mathbf{n}$ 's *structure function*. Here  $\delta \mathbf{n}$  is the perturbation of the index of refraction,  $\mathbf{X}$  is location in three dimensional space,  $t$  is time,  $\langle \dots \rangle$  is a spacetime average, and  $\sigma_n^2 \equiv \langle \delta \mathbf{n}^2 \rangle = C_n(0)$  is the variance of the fluctuations.

In Sec. 15.4.4 of Chap. 15, we shall show that, for strong and isotropic turbulence,  $D_n$  has the functional form  $D_n \propto \xi^{2/3}$  (where  $\xi \equiv |\boldsymbol{\xi}|$ ), with a multiplicative coefficient that we shall denote  $\mathcal{C}_n^2$  and that characterizes the strength of the perturbations:

$$D_n(\boldsymbol{\xi}) = \mathcal{C}_n^2 \xi^{2/3}. \quad (2)$$

The  $2/3$  power is called the *Kolmogorov spectrum* for the turbulence.

When light from a very distant star (a point source), directly overhead for simplicity, hits the earth's atmosphere, its phase fronts lie in horizontal planes, so the frequency- $\omega$  component of the electric field is  $\psi = e^{ikz}$ , where  $z$  increases downward. (Here we have factored out the field's overall amplitude.) When propagating through a thin layer of turbulent atmosphere, with thickness  $\delta h$ , the light acquires the phase fluctuation

$$\delta \varphi(\mathbf{x}, t) = k \int_z^{z+\delta h} \delta \mathbf{n}(\mathbf{x}, z, t) dz, \quad (3)$$

Here  $\mathbf{x}$  is the transverse, i.e. horizontal, location, and Eq. (3) follows from  $d\varphi = k dz$  with  $k = (\mathbf{n}/c)\omega$  and  $\mathbf{n} \simeq 1$ .

In Ex. 9.4 below, we derive some spacetime-averaged consequences of the phase fluctuations (3): (i) When the light emerges from the thin, turbulent layer, it has acquired a mean-square phase fluctuation on transverse lengthscale  $a$  given by

$$D_{\delta \varphi}(\mathbf{a}) \equiv \langle [\delta \varphi(\mathbf{x} + \mathbf{a}, t) - \delta \varphi(\mathbf{x}, t)]^2 \rangle = 2.91 \mathcal{C}_n^2 \delta h k^2 a^{5/3} \quad (4a)$$

[Eq. (2)], and a spacetime averaged complex visibility given by

$$\begin{aligned} \bar{\gamma}_\perp(k\mathbf{a}) &= \langle \psi(\mathbf{x}, t) \psi^*(\mathbf{x} + \mathbf{a}, t) \rangle = \langle \exp \{i [\delta \varphi(\mathbf{x}, t) - \delta \varphi(\mathbf{x} + \mathbf{a}, t)]\} \rangle \\ &= \exp \left[ -\frac{1}{2} D_{\delta \varphi}(a) \right] = \exp \left[ -1.455 \mathcal{C}_n^2 \delta h k^2 a^{5/3} \right]. \end{aligned} \quad (4b)$$

(ii) *Free propagation, including free-propagator diffraction effects which are important for long-distance propagation, preserves the spacetime-averaged complex visibility:  $d\bar{\gamma}_\perp/dz = 0$ .* (iii) Therefore, not surprisingly, when the turbulence is spread out vertically in some arbitrary manner, the net mean-square phase shift and time-averaged complex visibility observed on the ground are  $D_\varphi(\mathbf{a}) = 2.91 \left( \int \mathcal{C}_n^2(z) dz \right) k^2 a^{5/3}$ , and  $\bar{\gamma}_\perp(k\mathbf{a}) = \exp \left[ -\frac{1}{2} D_\varphi(a) \right]$ .

**Box 9.2 (Continued)**

It is conventional to introduce a transverse lengthscale  $r_o \equiv [0.423k^2 \int C_n^2(z)dz]^{-2/5}$  called the *Fried parameter*, in terms of which these  $D_\varphi$  and  $\bar{\gamma}_\perp$  are

$$D_\varphi(\mathbf{a}) = 6.88(a/r_o)^{5/3}, \quad (5a)$$

$$\bar{\gamma}_\perp(k\mathbf{a}) = \exp\left[-\frac{1}{2}D_\varphi(a)\right] = \exp[-3.44(a/r_o)^{5/3}]. \quad (5b)$$

This remarkably simple result provides opportunities to test the Kolmogorov power law: For light from a distant star, one can use a large telescope to measure  $\bar{\gamma}_\perp(ka)$ , and one can then plot  $\log \log \bar{\gamma}_\perp$  as a function of  $a$ . Equation (5b) predicts a slope 5/3 for this plot, and observations confirm that prediction.

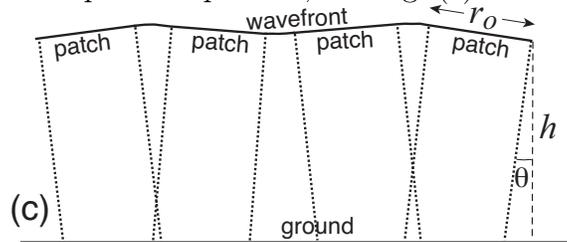
Notice that the Fried parameter  $r_o$  is the lengthscale on which the rms phase fluctuation  $\varphi_{\text{rms}} = \sqrt{D_\varphi(a=r_o)}$  is  $\sqrt{6.88} = 2.62$  radians; i.e.,  $r_o$  is the transverse length scale beyond which the turbulence-induced phase fluctuations are large compared to unity. These large random phase fluctuations drive  $\bar{\gamma}_\perp$  rapidly toward zero with increasing distance  $a$  [Eq. (5b)], i.e., they cause the light field to become spatially decorrelated with itself for distances  $a \gtrsim r_o$ . Therefore,  $r_o$  is (approximately) the time-averaged light field's spatial correlation length, on the ground. Moreover, since  $\bar{\gamma}_\perp$  is preserved under free propagation from the turbulent region to the ground,  $r_o$  must be the transverse correlation length of the light as it exits the turbulent region that produces the seeing. A correlated region at with transverse size  $r_o$ , is called an *isoplanatic patch*.

The observed time-averaged intensity  $\bar{I}(\boldsymbol{\alpha})$  from the point-source star is the Fourier transform of the complex visibility (5b); see Eq. (9.13b). This transform cannot be performed analytically, but a numerical computation gives the solid curve in Fig. (b) above, which agrees remarkably well with observations out to  $\sim 10^{-4}$  of the central intensity, where the Kolmogorov power law is expected to break down. Notice that the intensity distribution has a large-radius tail with far larger intensity than a Gaussian distribution (the dashed curve). This large-radius light is produced by large-angle diffraction, which is caused by very small-spatial-scale fluctuations (eddies) in the index of refraction.

Astronomers attribute to this time-averaged  $I(\boldsymbol{\alpha})$  an angular diameter  $\omega_o = 1.27\lambda/r_o$  (Ex. 9.4). For very good seeing conditions,  $r_o$  is about 20 cm and  $\omega_o$  is about 0.5 arcsec. Much more common is  $r_o \sim 10$  cm and  $\omega_o \sim 1$  arcsec.

**Speckel Pattern and Its Information**

The speckel pattern seen on short timescales,  $\lesssim 1/f_o \sim 0.01$  s, can be understood in terms of the turbulence's isoplanatic patches; see Fig. (c):



### Box 9.2 (Continued)

When the light field exits the turbulent region, at a height  $h \lesssim 1\text{km}$ , the isoplanatic patches on its wavefronts, with transverse size  $r_o$ , are planar to within roughly a reduced wavelength  $\lambda = 1/k = \lambda/2\pi$  (since the rms phase variation across a patch is just 2.62 radians). Each patch carries an image of the star, or whatever other object the astronomer is observing. The patch's light rays make an angle  $\theta \lesssim \omega_o = 1.27\lambda/r_o$  to the vertical. The patch's Fresnel length from the ground is  $r_F = \sqrt{\lambda h} \lesssim 2\text{ cm}$  (since  $\lambda \sim 0.5\mu\text{m}$  and  $h \lesssim 1\text{km}$ ). This is significantly smaller than the patch size  $r_o \sim 10$  to  $20\text{ cm}$ ; so there is little diffraction in the trip to ground. When these patches reach a large telescope (one with diameter  $D \gg r_o$ ), each is focused to produce an image of the object centered on the angular position  $\theta$  of its rays [dashed lines in Fig. (c)]. These images are the speckles of Fig. (a).

The speckle pattern [Fig. (a) above] varies rapidly because winds at high altitude sweep the isoplanatic patches through the star's light rays. For a wind speed  $u \sim 20\text{m s}^{-1}$ , the frequency of the pattern's fluctuations is  $f \sim u/r_o \sim 100\text{ Hz}$ , in agreement with observations.

To study the speckles and extract their information about the object's above-atmosphere intensity distribution  $I_o(\boldsymbol{\alpha})$ , one must observe them on timescales  $\lesssim 1/f \sim 10\text{ ms}$ . The first observations of this sort were the measurements of a few stellar diameters by Michelson and Pease, using the Michelson stellar interferometer (Sec. 9.2.7). The fringes that they saw were produced by the speckles, and because the phase of each speckle's fringes was random, the many speckles contributed incoherently to produce a far smaller net fringe visibility  $V$  than in the absence of atmospheric turbulence. Moreover, because the speckle pattern varied at  $f \sim 100\text{ Hz}$ , the net visibility and its phase also varied at  $f \sim 100\text{ Hz}$ . Fortunately, the human eye can discern things that vary this fast, so Michelson and Pease were able to see the fringes.

In the modern era of CCDs, fast electronics and powerful computers, a variety of more sophisticated techniques have been devised and implemented, for observing the speckles and extracting their information. These techniques fall into two classes: speckle image processing, and adaptive optics. Both are now (2013) in wide use by astronomers.

In *speckle image processing*, one makes optical measurements of the speckle pattern (sometimes with multi-pinhole masks), on timescales  $\lesssim 0.01\text{ s}$  for which the speckles are unchanging. One then uses optical or computational techniques to construct fourth-order or sixth-order correlations of the light field, e.g.  $\int \gamma_{\perp}[k(\mathbf{a} - \mathbf{a}')] \gamma_{\perp}^*(k\mathbf{a}') d^2a'$  (which is fourth-order in the field), from which a good approximation to the source's above-atmosphere intensity distribution  $I_o(\boldsymbol{\alpha})$  can be computed. A precursor to this was the Hanbury Brown & Twiss intensity interferometer, discussed in Sec. 9.6 below.

In *adaptive optics*, one focuses not on the speckles themselves, but on the turbulence-induced distortions of the wavefronts arriving the large telescope's mirror. One monitors those wavefront distortions with the aid of light from an artificial star (effectively a point source) that consists of fluorescing sodium atoms  $\sim 90\text{ km}$  above ground, which one

**Box 9.2 (Continued)**

excites with a laser beam. This artificial star must be within an angular distance  $\lesssim h/r_o \sim 3$  arcsec of the astronomical object one is observing (where  $h \sim 10$ km is the height of the highest turbulent layers that contribute significantly to the seeing). This  $\lesssim 3$  arcsec separation guarantees that light rays arriving at the same spot on the telescope mirror from the object and the artificial star will have traversed the same isoplanatic patches of turbulent atmosphere and thus have experienced the same phase delay and acquired the same wavefront distortions. One measures the wavefront distortions of the artificial star's light, and dynamically reshapes the telescope mirror so as to compensate for them, thereby removing the distortions not only from the artificial star's wavefronts, but also from the astronomical object's wavefronts. Thereby one converts the speckel pattern into the object's true intensity distribution  $I_o(\alpha)$ .

- (e) By combining (c) and (d), deduce Eqs. (5) for the mean-square phase fluctuations and spacetime-averaged visibility on the ground.
- (f) Perform a numerical Fourier transform of  $\gamma_{\perp}(k\mathbf{a})$  [Eq. (5b)] to get the time-averaged intensity distribution  $I(\alpha)$ . Construct a log-log plot of it, and compare with Fig. (b) of Box 9.2. What is  $r_o$  for the observational data shown in that figure?
- (g) Reexpress the turbulence-broadened image's spread angle  $\omega_o \equiv \{[4/\pi I(0)] \int I(\alpha) d\Omega_{\alpha}\}^{1/2}$  as an integral over  $\gamma_{\perp}$  and thence deduce that  $\omega_o = 1.27\lambda/r_o$ . Show that the Airy intensity distribution (8.18), for light that goes through a circular aperture with diameter  $D$ , also has  $\omega_o = 1.27D/\lambda$ .
- (g) It is conventional to define the angular diameter of the time-averaged image to be the diameter  $\omega_o$  of a constant-intensity disk, with the same central brightness  $I(0)$  as the image and the same total flux  $\int I(\alpha) d\Omega_{\alpha}$ ; i.e.,  $\pi(\omega_o/2)^2 I(0) = \int I(\alpha) d\Omega_{\alpha}$ . This  $\omega_o$  is called the image's *spread angle*. Reexpress  $\omega_o$  as an integral over  $\gamma_{\perp}$ . and thence deduce that  $\omega_o = 1.27\lambda/r_o$ . Show that the Airy intensity distribution (8.18), for light that goes through a circular aperture with diameter  $D$ , also has  $\omega_o = 1.27D/\lambda$ . (This same factor 1.27 is what motivated Fried's definition of  $r_o$ .) This implies that, *for a telescope with aperture  $D = r_o$ , turbulence-induced seeing produces the same degradation of the image as diffraction through the aperture. For smaller apertures,  $D < r_o$ , the aperture dominates the broadening of the time-averaged stellar image; for larger apertures,  $D > r_o$ , seeing is the dominant source of image broadening.*

\*\*\*\*\*

## 9.2.6 Temporal Coherence

In addition to the degree of spatial (or lateral) coherence, which measures the correlation of the field transverse to the direction of wave propagation, we can also measure the *degree*

of *temporal coherence*, also called the *degree of longitudinal coherence*. This describes the correlation at a given time at two points separated by a distance  $s$  along the direction of propagation. Equivalently, it measures the field sampled at a fixed position at two times differing by  $\tau = s/c$ . When (as in our discussion of spatial coherence) the waves are nearly monochromatic so the field arriving at the fixed position has the form  $\Psi = \psi(t)e^{-i\omega_o t}$ , then the degree of longitudinal coherence is complex and has a form completely analogous to the transverse case:

$$\gamma_{\parallel}(\tau) = \frac{\overline{\psi(t)\psi^*(t+\tau)}}{|\overline{\psi}|^2} \quad \text{for nearly monochromatic radiation} . \quad (9.15)$$

Here the average is over sufficiently long times  $t$  for the averaged value to settle down to an unchanging value.

When studying temporal coherence, one often wishes to deal with waves that contain a wide range of frequencies — e.g., the nearly Planckian (black-body) cosmic microwave radiation emerging from the very early universe (Ex. 9.6). In this case, one should not factor any  $e^{-i\omega_o t}$  out of the field  $\Psi$ , and one gains nothing by regarding  $\Psi(t)$  as complex, so the *longitudinal coherence*

$$\gamma_{\parallel}(\tau) = \frac{\overline{\Psi(t)\Psi(t+\tau)}}{\overline{\Psi}^2} \quad \text{for real } \Psi \text{ and broad-band radiation} \quad (9.16)$$

is also real. We shall use this real  $\gamma_{\parallel}$  throughout this subsection and the next. It obviously is the correlation function of  $\Psi$  [Eq. (6.19)] renormalized so  $\gamma_{\parallel}(0) = 1$ .

As  $\tau$  is increased,  $\gamma_{\parallel}$  typically remains near unity until some critical value  $\tau_c$  is reached, and then begins to fall off toward zero. The critical value  $\tau_c$ , the longest time over which the field is strongly coherent, is the *coherence time*, of which we have already spoken: If the wave is roughly monochromatic so  $\Psi(t) \propto \cos[\omega_o t + \delta\varphi(t)]$ , with  $\omega_o$  fixed and the phase  $\delta\varphi$  randomly varying in time, then it should be clear that the mean time for  $\delta\varphi$  to change by an amount of order unity is the coherence time  $\tau_c$  at which  $\gamma_{\parallel}$  begins to fall significantly.

The uncertainty principle dictates that a field with coherence time  $\tau_c$ , when Fourier analyzed in time, must contain significant power over a bandwidth  $\Delta f = \Delta\omega/2\pi \sim 1/\tau_c$ . Correspondingly, if we define the field's *longitudinal coherence length* by

$$l_{\parallel} \equiv c\tau_c , \quad (9.17)$$

then  $l_{\parallel}$  for broad-band radiation will be only a few times the peak wavelength, but for a narrow spectral line of width  $\Delta\lambda$ , it will be  $\lambda^2/\Delta\lambda$ .

These relations between the coherence time or longitudinal coherence length and the field's spectral energy flux are order-of-magnitude consequences not only of the uncertainty relation, but also of the temporal analog of the van Cittert-Zernike Theorem. That analog is just the Wiener-Khintchine Theorem in disguise, and it can be derived by the same methods as we used in the transverse spatial domain. It says that the degree of lateral coherence  $\gamma_{\perp}$  is replaced by the degree of temporal coherence  $\gamma_{\parallel}$ , and the angular intensity distribution  $I(\alpha)$  (distribution of energy over angle) is replaced by the field's spectral energy flux  $F_{\omega}(\omega)$

(the energy crossing a unit area per unit time and per unit angular frequency  $\omega$ )—which is also called its *spectrum*.<sup>4</sup> The theorem takes the explicit form

$$\gamma_{\parallel}(\tau) = \frac{\int_{-\infty}^{\infty} d\omega F_{\omega}(\omega) e^{i\omega\tau}}{F_S} = \frac{2 \int_0^{\infty} d\omega F_{\omega}(\omega) \cos \omega\tau}{F_s} \quad \text{for real } \Psi(t), \text{ valid for broad-band radiation} \quad (9.18a)$$

and

$$F_{\omega}(\omega) = F_S \int_{-\infty}^{\infty} \frac{d\tau}{2\pi} \gamma_{\parallel}(\tau) e^{-i\omega\tau} = 2F_s \int_0^{\infty} \frac{d\tau}{2\pi} \gamma_{\parallel}(\tau) \cos \omega\tau. \quad (9.18b)$$

[Here the normalization of our Fourier transform and the sign of its exponential are those conventionally used in optics, and differ from those used in the theory of random processes (Chap. 6). Also, because we have chosen  $\Psi$  to be real,  $F_{\omega}(-\omega) = F_{\omega}(+\omega)$  and  $\gamma_{\parallel}(-\tau) = \gamma_{\parallel}(+\tau)$ .] One can measure  $\gamma_{\parallel}$  by combining the radiation from two points displaced longitudinally to produce interference fringes just as we did in measuring spatial coherence. This type of interference is sometimes called *interference by division of the amplitude*, in contrast with “interference by division of the wave front” for a Young’s-slit-type measurement of lateral spatial coherence (next to the last paragraph of Sec. 9.2.1).

\*\*\*\*\*

## EXERCISES

**Exercise 9.5** *Problem: Longitudinal coherence of radio waves*

An FM radio station has a carrier frequency of 91.3 MHz and transmits heavy metal rock music in frequency modulated side bands of the carrier. Estimate the coherence length of the radiation.

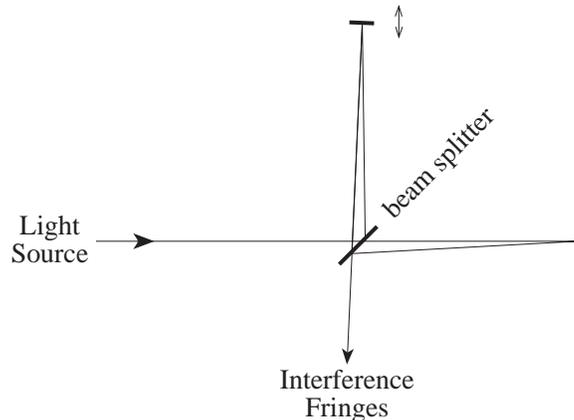
\*\*\*\*\*

### 9.2.7 Michelson Interferometer and Fourier-Transform Spectroscopy

The classic instrument for measuring the degree of longitudinal coherence is the Michelson interferometer of Fig. 9.3 (not to be confused with the Michelson stellar interferometer). In the simplest version, incident light (e.g. in the form of a Gaussian beam; Sec. 8.5.5) is split by a beam splitter into two beams, which are reflected off different plane mirrors and then recombined. The relative positions of the mirrors are adjustable so that the two light paths can have slightly different lengths. (An early version of this instrument was used in the famous Michelson-Morley experiment.) There are two ways to view the fringes. One way is to tilt one of the reflecting mirrors slightly so there is a range of path lengths in one

---

<sup>4</sup>Note that the spectral energy flux (spectrum) is simply related to the spectral density of the field: If the field  $\Psi$  is so normalized that the energy density is  $U = \beta \overline{\Psi_{,t} \Psi_{,t}}$  with  $\beta$  some constant, then  $F_{\omega}(\omega) = \beta c / (2\pi) S_{\Psi}(f)$ , with  $f = \omega / 2\pi$ .



**Fig. 9.3:** Michelson Interferometer.

of the arms. Light and dark interference bands (fringes) can then be seen across the circular cross section of the recombined beam. The second method is conceptually more direct but requires aligning the mirrors sufficiently accurately so the phase fronts of the two beams are parallel after recombination and the recombined beam has no banded structure. The end mirror in one arm of the interferometer is then slowly moved backward or forward, and as it moves, the recombined light slowly changes from dark to light to dark and so on.

It is interesting to interpret this second method in terms of the Doppler shift. One beam of light undergoes a Doppler shift on reflection off the moving mirror. There is then a beat wave produced when it is recombined with the unshifted radiation of the other beam.

Whichever method is used (tilted mirror or longitudinal motion of mirror), the visibility  $\gamma_{\parallel}$  of the interference fringes measures the beam's degree of longitudinal coherence, which is related to the spectral energy flux (spectrum)  $F_{\omega}$  by Eqs. (9.18).

Let us give an example. Suppose we observe a spectral line with rest angular frequency  $\omega_0$  that is broadened by random thermal motions of the emitting atoms so the line profile is

$$F_{\omega} \propto \exp\left(-\frac{(\omega_0 - \omega)^2}{2(\Delta\omega)^2}\right). \quad (9.19a)$$

The width of the line is given by the formula for the Doppler shift,  $\Delta\omega \sim \omega_0(k_B T/mc^2)^{1/2}$ , where  $T$  is the temperature of the emitting atoms and  $m$  is their mass. (We ignore other sources of line broadening, e.g. natural broadening and pressure broadening, which actually dominate under normal conditions.) For example, for hydrogen at  $T = 10^3\text{K}$ , the Doppler-broadened line width is  $\Delta\omega \sim 10^{-5}\omega_0$ .

By Fourier transforming this line profile, using the well known result that the Fourier transform of a Gaussian is another Gaussian, and invoking the fundamental relations (9.18) between the spectrum and temporal coherence, we obtain

$$\gamma_{\parallel}(\tau) = \exp\left(-\frac{\tau^2(\Delta\omega)^2}{2}\right) \cos \omega_0 \tau. \quad (9.19b)$$

If we had used the nearly monochromatic formalism with the field written as  $\Psi = \psi(t)e^{-i\omega_0 t}$ ,

then we would have obtained

$$\gamma_{\parallel}(\tau) = \exp\left(-\frac{\tau^2(\Delta\omega)^2}{2}\right) e^{i\omega_0\tau}, \quad (9.19c)$$

the real part of which is our broad-band formalism's  $\gamma_{\parallel}$ . In either case,  $\gamma_{\parallel}$  oscillates with angular frequency  $\omega_0$ , and the amplitude of this oscillation is the *fringe visibility*  $V$ :

$$V = \exp\left(-\frac{\tau^2(\Delta\omega)^2}{2}\right). \quad (9.19d)$$

The variation  $V(\tau)$  of this visibility with lag time  $\tau$  is sometimes called an *interferogram*. For time lags  $\tau \ll (\Delta\omega)^{-1}$ , the line appears to be monochromatic and fringes with unit visibility should be seen. However, for lags  $\tau \gtrsim (\Delta\omega)^{-1}$ , the fringe visibility will decrease exponentially with  $\tau^2$ . In our Doppler-broadened hydrogen-line example with  $\Delta\omega \sim 10^{-5}\omega_0$ , the rest angular frequency is  $\omega_0 \sim 3 \times 10^{15} \text{ rad s}^{-1}$ , so the longitudinal coherence length is  $l_{\parallel} = c\tau_c \sim 10\text{mm}$ . No fringes will be seen when the radiation is combined from points separated by much more than this distance.

This procedure is an example of *Fourier transform spectroscopy*, in which, by measuring the degree of temporal coherence  $\gamma_{\parallel}(\tau)$  and then Fourier transforming it [Eq. (9.18)], one infers the shape of the radiation's spectrum, or in this case, the width of a specific spectral line.

When (as in Ex. 9.6 below) the waves are very broad band, the degree of longitudinal coherence  $\gamma_{\parallel}(\tau)$  will not have the form of a sinusoidal oscillation (regular fringes) with slowly varying amplitude (visibility). Nevertheless, the broad-band van Cittert-Zernike theorem (9.18) still guarantees that the spectrum (spectral energy flux) will be the Fourier transform of the coherence  $\gamma_{\parallel}(\tau)$ , which can be measured by a Michelson interferometer.

\*\*\*\*\*

## EXERCISES

**Exercise 9.6** *Problem: COBE Measurement of the Cosmic Microwave Background Radiation*

An example of a Michelson interferometer is the Far Infrared Absolute Spectrophotometer (FIRAS) carried by the Cosmic Background Explorer Satellite (COBE). COBE studied the spectrum and anisotropies of the *cosmic microwave background radiation* (CMB) that emerged from the very early, hot phase of our universe's expansion (Chap. 28). One of the goals of the COBE mission was to see if the CMB spectrum really had the shape of 2.7K black body (Planckian) radiation, or if it was highly distorted as some measurements made on rocket flights had suggested. COBE's spectrophotometer used Fourier transform spectroscopy to meet this goal: it compared accurately the degree of longitudinal coherence  $\gamma_{\parallel}$  of the CMB radiation with that of a calibrated source on board the spacecraft, which was known to be a black body at about 2.7K. The comparison was made by alternately feeding radiation from the microwave background and radiation from the calibrated source into the same Michelson interferometer and comparing their fringe spacings. The result (Mather et.

al. 1994) was that the background radiation has a spectrum that is Planckian with temperature  $2.726 \pm 0.010\text{K}$  over the wavelength range 0.5–5 mm, in agreement with simple cosmological theory that we shall explore in the last chapter of this book.

- (a) Suppose that the CMB had had a Wien spectrum  $F_\omega \propto |\omega|^3 \exp(-\hbar|\omega|/kT)$  where  $T = 2.74\text{K}$ . Show that the *visibility* of the fringes would have been

$$V = |\gamma_{||}| \propto \frac{|s^4 - 6s_0^2s^2 + s_0^4|}{(s^2 + s_0^2)^4} \quad (9.20)$$

where  $s = c\tau$  is longitudinal distance, and calculate a numerical value for  $s_0$ .

- (b) Compute the interferogram  $V(\tau)$  for a Planck function either analytically (perhaps with the help of a computer) or numerically using a Fast Fourier Transform. Compare graphically the interferogram for the Wien and Planck spectra.

\*\*\*\*\*

## 9.2.8 Degree of Coherence; Relation to Theory of Random Processes

Having separately discussed spatial and temporal coherence, we now can easily perform a final generalization and define the full degree of coherence of the radiation field between two points separated both laterally by a vector  $\mathbf{a}$  and longitudinally by a distance  $s$ , or equivalently by a time  $\tau = s/c$ . If we restrict ourselves to nearly monochromatic waves and use the complex formalism so the waves are written as  $\Psi = e^{i(kz - \omega t)}\psi(\mathbf{x}, t)$  [Eq. (9.12a)], then

$$\gamma_{12}(\mathbf{ka}, \tau) \equiv \frac{\overline{\psi(\mathbf{x}_1, t)\psi^*(\mathbf{x}_1 + \mathbf{a}, t + \tau)}}{[\overline{|\psi(\mathbf{x}_1, t)|^2} \overline{|\psi(\mathbf{x}_1 + \mathbf{a}, t)|^2}]^{1/2}} = \frac{\overline{\psi(\mathbf{x}_1, t)\psi^*(\mathbf{x}_1 + \mathbf{a}, t + \tau)}}{\overline{|\psi|^2}}. \quad (9.21)$$

In the denominator of the second expression we have used the fact that, because the source is far away,  $\overline{|\psi|^2}$  is independent of the spatial location at which it is evaluated, in the region of interest. Consistent with the definition (9.21), we can define a *volume of coherence*  $\mathcal{V}_c$  as the product of the longitudinal coherence length  $l_{||} = c\tau_c$  and the square of the transverse coherence length  $l_{\perp}^2$ :  $\mathcal{V}_c = l_{\perp}^2 c\tau_c$ .

The three-dimensional version of the van Cittert-Zernike theorem relates the complex degree of coherence (9.21) to the radiation's *spectral intensity*,  $I_\omega(\boldsymbol{\alpha}, \omega)$ , i.e. to the energy crossing a unit area per unit time per unit solid angle and per unit angular frequency (energy “per unit everything”). (Since the frequency  $\nu = f$  and the angular frequency  $\omega$  are related by  $\omega = 2\pi\nu$ , the spectral intensity  $I_\omega$  of this chapter and that  $I_\nu$  of Chap. 3 are related by  $I_\nu = 2\pi I_\omega$ .) The *three-dimensional van Cittert-Zernike theorem* states that

$$\boxed{\gamma_{12}(\mathbf{ka}, \tau) = \frac{\int d\Omega_\alpha d\omega I_\omega(\boldsymbol{\alpha}, \omega) e^{i(\mathbf{ka} \cdot \boldsymbol{\alpha} + \omega\tau)} }{F_S}}, \quad (9.22a)$$

and

$$I_\omega(\boldsymbol{\alpha}, \omega) = F_S \int \frac{d\tau d^2ka}{(2\pi)^3} \gamma_{12}(\mathbf{k}\mathbf{a}, \tau) e^{-i(\mathbf{k}\mathbf{a}\cdot\boldsymbol{\alpha} + \omega\tau)}. \quad (9.22b)$$

There obviously must be an intimate relationship between the theory of random processes, as developed in Chap. 6, and the theory of a wave's coherence, as we have developed it in this section, Sec. 9.2. That relationship is explained in Ex. 9.8. Most especially, it is shown that *the van Cittert-Zernike theorem is nothing but the wave's Wiener-Khintchine theorem in disguise.*

\*\*\*\*\*

## EXERCISES

### Exercise 9.7 *Problem: Decomposition of Degree of Coherence*

We have defined the degree of coherence  $\gamma_{12}(\mathbf{a}, \tau)$  for two points in the radiation field separated laterally by a distance  $\mathbf{a}$  and longitudinally by a time  $\tau$ . Under what conditions will this be given by the product of the spatial and temporal degrees of coherence?

$$\gamma_{12}(\mathbf{a}, \tau) = \gamma_\perp(\mathbf{a})\gamma_\parallel(\tau) \quad (9.23)$$

### Exercise 9.8 \*\*\* *Example: Complex Random Processes and the van Cittert-Zernike Theorem*

In Chap. 6 we developed the theory of real-valued random processes that vary randomly with time  $t$ , i.e. that are defined on a one-dimensional space in which  $t$  is a coordinate. Here we shall generalize a few elements of that theory to a complex-valued random process  $\Phi(\mathbf{x})$  defined on a (Euclidean) space with  $n$  dimensions. We assume the process to be stationary and to have vanishing mean (cf. Chap. 6 for definitions). For  $\Phi(\mathbf{x})$  we define a complex-valued correlation function by

$$C_\Phi(\boldsymbol{\xi}) \equiv \overline{\Phi(\mathbf{x})\Phi^*(\mathbf{x} + \boldsymbol{\xi})} \quad (9.24a)$$

(where the  $*$  denotes complex conjugation) and a real-valued spectral density by

$$S_\Phi(\mathbf{k}) = \lim_{L \rightarrow \infty} \frac{1}{L^n} |\tilde{\Phi}_L(\mathbf{k})|^2. \quad (9.24b)$$

Here  $\Phi_L$  is  $\Phi$  confined to a box of side  $L$  (i.e. set to zero outside that box), and the tilde denotes a Fourier transform defined using the conventions of Chap. 6:

$$\tilde{\Phi}_L(\mathbf{k}) = \int \Phi_L(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^n x, \quad \Phi_L(\mathbf{x}) = \int \tilde{\Phi}_L(\mathbf{k}) e^{+i\mathbf{k}\cdot\mathbf{x}} \frac{d^n k}{(2\pi)^n}. \quad (9.25)$$

Because  $\Phi$  is complex rather than real,  $C_\Phi(\boldsymbol{\xi})$  is complex; and as we shall see below, its complexity implies that [although  $S_\Phi(\mathbf{k})$  is real],  $S_\Phi(-\mathbf{k}) \neq S_\Phi(\mathbf{k})$ . This fact prevents us from folding negative  $\mathbf{k}$  into positive  $\mathbf{k}$  and thereby making  $S_\Phi(\mathbf{k})$  into a "single-sided" spectral density as we did for real random processes in Chap. 6. In this complex case we must distinguish  $-\mathbf{k}$  from  $+\mathbf{k}$  and similarly  $-\boldsymbol{\xi}$  from  $+\boldsymbol{\xi}$ .

(a) The complex Wiener-Khintchine theorem [analog of Eq.(6.29)] says that

$$S_{\Phi}(\mathbf{k}) = \int C_{\Phi}(\boldsymbol{\xi}) e^{+i\mathbf{k}\cdot\boldsymbol{\xi}} d^n \xi, \quad (9.26a)$$

$$C_{\Phi}(\boldsymbol{\xi}) = \int S_{\Phi}(\mathbf{k}) e^{-i\mathbf{k}\cdot\boldsymbol{\xi}} \frac{d^n k}{(2\pi)^n}. \quad (9.26b)$$

Derive these relations. [Hint: use Parseval's theorem in the form  $\int A(\mathbf{x})B^*(\mathbf{x})d^n x = \int \tilde{A}(\mathbf{k})\tilde{B}^*(\mathbf{k})d^n k/(2\pi)^n$  with  $A(\mathbf{x}) = \Phi_L(\mathbf{x})$  and  $B(\mathbf{x}) = \Phi_L(\mathbf{x} + \boldsymbol{\xi})$ , and then take the limit as  $L \rightarrow \infty$ .] Because  $S_{\Phi}(\mathbf{k})$  is real, this Wiener-Khintchine theorem implies that  $C_{\Phi}(-\boldsymbol{\xi}) = C_{\Phi}^*(\boldsymbol{\xi})$ . Show that this is so directly from the definition (9.24a) of  $C_{\Phi}(\boldsymbol{\xi})$ . Because  $C_{\Phi}(\boldsymbol{\xi})$  is complex, the Wiener-Khintchine theorem implies that  $S_{\Phi}(\mathbf{k}) \neq S_{\Phi}(-\mathbf{k})$ .

(b) Let  $\psi(\mathbf{x}, t)$  be the complex-valued wave field defined in Eq. (9.12a), and restrict  $\mathbf{x}$  to range only over the two transverse dimensions so  $\psi$  is defined on a 3-dimensional space. Define  $\Phi(\mathbf{x}, t) \equiv \psi(\mathbf{x}, t)/[|\psi(\mathbf{x}, t)|^2]^{1/2}$ . Show that

$$C_{\Phi}(\mathbf{a}, \tau) = \gamma_{12}(k\mathbf{a}, \tau), \quad S_{\Phi}(-\boldsymbol{\alpha}k, -\omega) = \text{const} \times \frac{I_{\omega}(\boldsymbol{\alpha}, \omega)}{F_S}, \quad (9.27)$$

and that the complex Wiener-Khintchine theorem (9.26) is the van Cittert-Zernike theorem (9.22). (Note: the minus signs in  $S_{\Phi}$  result from the difference in Fourier transform conventions between the theory of random processes [Eq. (9.25) above and Chap. 6] and the theory of optical coherence [this chapter]. Evaluate the constant in Eq. (9.27).

\*\*\*\*\*

### 9.3 Radio Telescopes

The technique pioneered by Michelson for measuring the angular sizes of stars at visual wavelengths has been applied with great effect in radio astronomy. A modern radio telescope is a large, steerable surface that reflects radio waves onto a “feed”, where the fluctuating electric field in the radio wave creates a very small electric voltage that can subsequently be amplified and measured electronically. A large telescope has a diameter  $D \sim 100\text{m}$  and a typical observing wavelength might be  $\lambda \sim 6\text{cm}$ . This implies an angular resolution  $\theta_A \sim \lambda/D \sim 2$  arc minutes [Eq. (8.18) and subsequent discussion]. However, many of the most interesting cosmic sources are much smaller than this. In order to achieve much better angular resolution, the technique of radio interferometry was developed in the 1960s and 70s; and the analogous optical interferometry is currently (2010s) under rapid development.<sup>5</sup>

<sup>5</sup>Optical interferometry was not possible until optical technology became good enough to monitor the phase of light, as well as its amplitude, at separate locations and then produce interference.

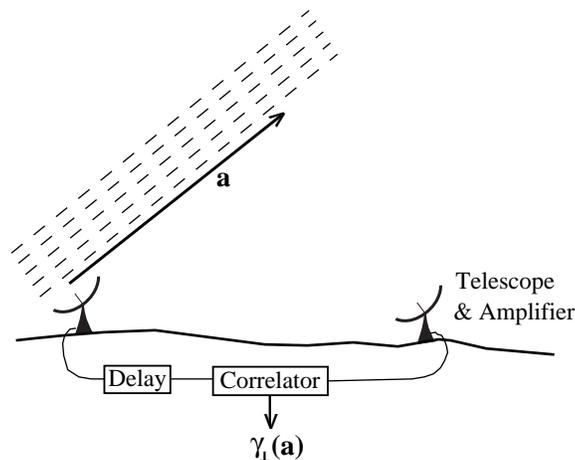
### 9.3.1 Two-Element Radio Interferometer

If we have two radio telescopes, then we can think of them as two Young's slits, and we can link them using a combination of waveguides and electric cables as shown in Fig. 9.4. When they are both pointed at a source, they both measure the electric field in radio waves from that source. We combine their signals by narrow-band filtering their voltages to make them nearly monochromatic and then either add the filtered voltages and measure the power, or multiply the two voltages directly. In either case a measurement of the degree of coherence, Eq. (9.10) can be achieved. (If the source is not vertically above the two telescopes, one obtains some non-lateral component of the full degree of coherence  $\gamma_{12}(\mathbf{a}, \tau)$ . However, by introducing a time delay into one of the signals, as in Fig. 9.4, one can measure the degree of lateral coherence  $\gamma_{\perp}(\mathbf{a})$ , which is what the astronomer usually needs.)

The objective is usually to produce an image of the radio waves' source. This is achieved by Fourier inverting the lateral degree of coherence  $\gamma_{\perp}(\mathbf{a})$  [Eq. (9.13b)], which must therefore be measured for a variety of values of the relative separation vector  $\mathbf{a}$  of the telescopes perpendicular to the direction of the source. As the earth rotates, the separation vector will trace out half an ellipse in the two-dimensional  $\mathbf{a}$  plane every twelve hours. [The source intensity is a real quantity and so we can use Eq. (9.13b) to deduce that  $\gamma_{\perp}(-\mathbf{a}) = \gamma_{\perp}^*(\mathbf{a})$ , which gives the other half of the ellipse.] By changing the spacing between the two telescopes daily and collecting data for a number of days, the degree of coherence can be well sampled. This technique is known as *Earth-Rotation Aperture Synthesis* because the telescopes are being made to behave like a giant telescope, as big as their maximum separation, with the aid of the earth's rotation.

### 9.3.2 Multiple Element Radio Interferometer

In practice, a modern radio interferometer has many more than two telescopes. The Very Large Array (VLA) in New Mexico (USA) has 27 individual telescopes arranged in a Y pattern and operating simultaneously. The degree of coherence can thus be measured simultane-



**Fig. 9.4:** Two element radio interferometer.

ously over  $27 \times 26/2 = 351$  different relative separations. The results of these measurements can then be interpolated to give values of  $\gamma_{\perp}(\mathbf{a})$  on a regular grid of points (usually  $2^N \times 2^N$  for some integer  $N$ ). This is then suitable for applying the Fast Fourier Transform algorithm to infer the source structure  $I(\alpha)$ .

### 9.3.3 Closure Phase

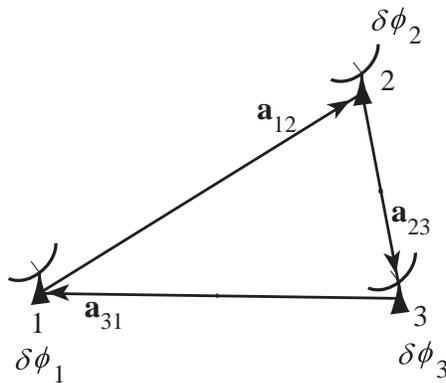
Among the many technical complications of interferometry is one which brings out an interesting point about Fourier methods. It is usually much easier to measure the modulus than the phase of the complex degree of coherence. This is partly because it is hard to introduce the necessary delays in the electronics accurately enough to know where the zero of the fringe pattern should be located and partly because unknown, fluctuating phase delays are introduced into the phase of the field as the wave propagates through the upper atmosphere and ionosphere. (This is a radio variant of the problem of “seeing” for optical telescopes, cf. Ex. 8.10, and it also plagues the Michelson stellar interferometer.) It might therefore be thought that we would have to make do with just the modulus of the degree of coherence, i.e. the fringe visibility, to perform the Fourier inversion for the source structure. This is not so.

Consider a three element interferometer measuring fields  $\psi_1, \psi_2, \psi_3$  and suppose that at each telescope there are unknown phase errors,  $\delta\varphi_1, \delta\varphi_2, \delta\varphi_3$ ; cf. Fig. 9.5. For baseline  $\mathbf{a}_{12}$ , we measure the degree of coherence  $\gamma_{\perp 12} \propto \overline{\psi_1 \psi_2^*}$ , a complex number with phase  $\Phi_{12} = \varphi_{12} + \delta\varphi_1 - \delta\varphi_2$ , where  $\varphi_{12}$  is the phase of  $\gamma_{\perp 12}$  in the absence of phase errors. If we also measure the degrees of coherence for the other two pairs of telescopes in the triangle and derive their phases  $\Phi_{23}, \Phi_{31}$ , we can then calculate the quantity

$$\begin{aligned} C_{123} &= \Phi_{12} + \Phi_{23} + \Phi_{31} \\ &= \varphi_{12} + \varphi_{23} + \varphi_{31}, \end{aligned} \quad (9.28)$$

from which the phase errors cancel out.

The quantity  $C_{123}$ , known as the *closure phase*, can be measured with high accuracy. In the VLA, there are  $27 \times 26 \times 25/6 = 2925$  such closure phases, and they can all be measured



**Fig. 9.5:** Closure phase measurement using a triangle of telescopes.

with considerable redundancy. Although absolute phase information cannot be recovered, 93 per cent of the telescopes' relative phases can be inferred in this manner and used to construct an image far superior to what one would get without any phase information.

### 9.3.4 Angular Resolution

When the telescope spacings are well sampled and the source is bright enough to carry out these image processing techniques, an interferometer can have an angular resolving power approaching that of an equivalent filled aperture as large as the maximum telescope spacing. For the VLA this is 35km, giving an angular resolution of a fraction of a second of arc at 6cm wavelength, which is 350 times better than the resolution of a single 100m telescope.

Even greater angular resolution is achieved in a technique known as Very Long Baseline Interferometry (VLBI). Here the telescopes can be located on different continents and instead of linking them directly, the oscillating field amplitudes  $\psi(t)$  are stored electronically and then combined digitally long after the observation, to compute the complex degree of coherence and thence the source structure  $I(\boldsymbol{\alpha})$ . In this way angular resolutions over 300 times better than those achievable by the VLA have been obtained. Structure smaller than a milliarcsecond corresponding to a few light years at cosmological distances can be measured in this manner.

This technique will be used in the international Square Kilometer Array (SKA), consisting of thousands of telescopes spread over linear distances of several thousand kilometers, with high-telescope-density central cores in South Africa and Australia. The total area of the telescope dishes (total "collecting area") will be about one square kilometer; hence the name SKA.

\*\*\*\*\*

## EXERCISES

### Exercise 9.9 Example: Radio Interferometry from Space

The longest radio-telescope separation available in 2012 is that between telescopes on the earth's surface and a 10-m diameter radio telescope in the Russian RadioAstron satellite, which was launched into a highly elliptical orbit around Earth in summer 2011, with perigee  $\sim 10,000$  km (1.6 Earth radii) and apogee  $\sim 350,000$  km (55 Earth radii).

- (a) Radio astronomers conventionally describe the spectral intensity  $I_\omega(\boldsymbol{\alpha}, \omega)$  of a source in terms of its brightness temperature. This is the temperature  $T_b(\omega)$  that a black body would have to have in order to emit, in the Rayleigh-Jeans (low-frequency) end of its spectrum, the same spectral intensity as the source. Show that for a single (linear or circular) polarization, if the solid angle subtended by a source is  $\Delta\Omega$  and the *spectral energy flux* measured from the source is  $F_\omega \equiv \int I_\omega d\Omega = I_\omega \Delta\Omega$ , then the brightness temperature is

$$T_b = \frac{(2\pi)^3 c^2 I_\omega}{k_B \omega^2} = \frac{(2\pi)^3 c^2 F_\omega}{k_B \omega^2 \Delta\Omega}, \quad (9.29)$$

where  $k_B$  is Boltzmann's constant.

- (b) The brightest quasars emit radio spectral fluxes of about  $F_\omega = 10^{-25} \text{W m}^{-2} \text{Hz}^{-1}$ , independent of frequency. The smaller is such a quasar, the larger will be its brightness temperature. Thus, one can characterize the smallest sources that a radio telescope system can resolve by the highest brightness temperatures it can measure. Show that the maximum brightness temperature measurable by the earth-to-orbit RadioAstron interferometer is independent of the frequency at which the observation is made, and estimate its numerical value.

\*\*\*\*\*

## 9.4 Etalons and Fabry-Perot Interferometers

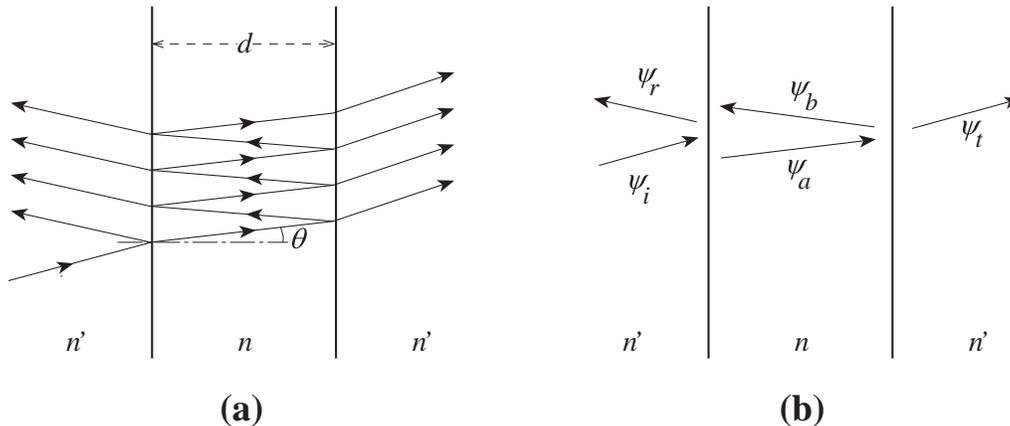
We have shown how a Michelson interferometer (Fig. 9.3) can be used as a Fourier-transform spectrometer: one measures the complex fringe visibility as a function of the two arms' optical path difference and then takes the visibility's Fourier transform to obtain the spectrum of the radiation. The inverse process is also powerful: One can drive a Michelson interferometer with radiation with a known, steady spectrum (usually close to monochromatic), and look for time variations of the positions of its fringes caused by changes in the relative optical path lengths of the interferometer's two arms. This was the philosophy of the famous Michelson-Morley experiment to search for ether drift, and it is also the underlying principle of a laser interferometer ("interferometric") gravitational-wave detector.

To reach the sensitivity required for gravitational-wave detection, one must modify the Michelson interferometer by making the light travel back and forth in each arm many times, thereby amplifying the phase shift caused by changes in the arm lengths. This is achieved by converting each arm into a Fabry-Perot interferometer. In this section we shall study Fabry-Perot interferometers and some of their other applications, and in the next section we shall explore their use in gravitational-wave detection.

### 9.4.1 Multiple Beam Interferometry; Etalons

Fabry-Perot interferometry is based on trapping monochromatic light between two highly reflecting surfaces. To understand such trapping, let us consider the concrete situation where the reflecting surfaces are flat and parallel to each other, and the transparent medium between the surfaces has one index of refraction  $\mathbf{n}$ , while the medium outside the surfaces has another index  $\mathbf{n}'$  (Fig. 9.6). Such a device is sometimes called an *etalon*. One example is a glass slab in air ( $\mathbf{n} \simeq 1.5, \mathbf{n}' \simeq 1$ ); another is a vacuum maintained between two glass mirrors ( $\mathbf{n} = 1, \mathbf{n}' \simeq 1.5$ ). For concreteness, we shall discuss the slab case, though all our formulas are valid equally well for vacuum between mirrors or for any other etalon.

Suppose that a monochromatic plane wave (i.e. parallel rays) with angular frequency  $\omega$  is incident on one of the slab's reflecting surfaces, where it is partially reflected and partially transmitted with refraction. The transmitted wave will propagate through to the second surface, where it will be partially reflected and partially transmitted. The reflected portion



**Fig. 9.6:** Multiple beam interferometry using a type of Fabry-Perot etalon.

will return to the first surface, where it too will be split, and so on (Fig. 9.6a). The resulting total fields in and outside the slab can be computed by summing the series of sequential reflections and transmissions (Ex. 9.12). Alternatively, they can be computed as follows:

We shall assume, for pedagogical simplicity, that there is translational invariance along the slab (i.e. the slab and incoming wave are perfectly planar). Then the series, if summed, would lead to the five waves shown in Fig. 9.6b: an incident wave ( $\psi_i$ ), a reflected wave ( $\psi_r$ ), a transmitted wave ( $\psi_t$ ), and two internal waves ( $\psi_a$  and  $\psi_b$ ).

We introduce amplitude reflection and transmission coefficients, denoted  $\mathbf{r}$  and  $\mathbf{t}$ , for waves incident upon the slab surface from outside. Likewise, we introduce coefficients  $\mathbf{r}'$ ,  $\mathbf{t}'$  for waves incident upon the slab from inside. These coefficients are functions of the angles of incidence and the light's polarization. They can be computed using electromagnetic theory (e.g. Sec. 4.6.2 of Hecht 2002), but this will not concern us here.

Armed with these definitions, we can express the reflected and transmitted waves at the first surface (location  $A$  in Fig. 9.7) in the form

$$\begin{aligned}\psi_r &= \mathbf{r}\psi_i + \mathbf{t}'\psi_b, \\ \psi_a &= \mathbf{t}\psi_i + \mathbf{r}'\psi_b,\end{aligned}\tag{9.30a}$$

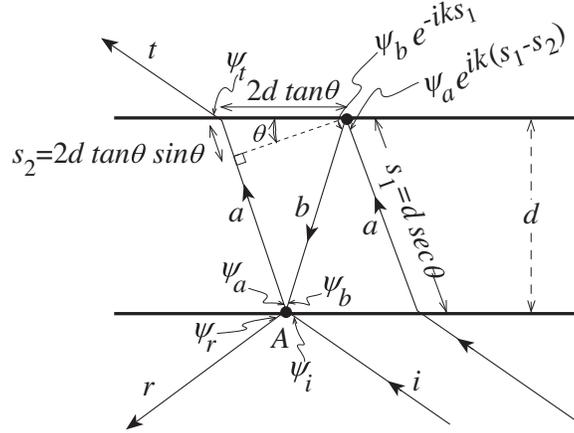
where  $\psi_i$ ,  $\psi_a$ ,  $\psi_b$ , and  $\psi_r$  are the values of  $\psi$  at  $A$  for waves impinging on or leaving the surface along the paths  $i$ ,  $a$ ,  $b$ , and  $r$  depicted in Fig. 9.7. Simple geometry shows that the waves at the second surface are as depicted in Fig. 9.7; and correspondingly, the relationships between the ingoing and outgoing waves there are

$$\begin{aligned}\psi_b e^{-iks_1} &= \mathbf{r}'\psi_a e^{ik(s_1-s_2)}, \\ \psi_t &= \mathbf{t}'\psi_a e^{iks_1},\end{aligned}\tag{9.30b}$$

where  $k = n\omega/c$  is the wave number in the slab and (as is shown in the figure)

$$s_1 = d \sec \theta, \quad s_2 = 2d \tan \theta \sin \theta,\tag{9.30c}$$

with  $d$  the thickness of the slab and  $\theta$  the angle that the wave fronts inside the slab make to the slab's faces.



**Fig. 9.7:** Construction for calculating the phase differences across the slab for the two internal waves in an etalon.

In solving Eqs. (9.30) for the net transmitted and reflected waves  $\psi_t$  and  $\psi_r$  in terms of the incident wave  $\psi_i$ , we shall need *reciprocity relations* between the reflection and transmission coefficients  $\mathbf{r}$ ,  $\mathbf{t}$  for waves that hit the reflecting surfaces from one side, and those  $\mathbf{r}'$ ,  $\mathbf{t}'$  for waves from the other side. These reciprocity relations are analyzed quite generally in Ex. 9.10. To derive the reciprocity relations in our case of sharp boundaries between homogeneous media, consider the limit in which the slab thickness  $d \rightarrow 0$ . This is allowed because the wave equation is linear and the solution for one surface can be superposed on that for the other surface. In this limit  $s_1 = s_2 = 0$  and the slab must become transparent so

$$\psi_r = 0, \quad \psi_t = \psi_i. \quad (9.31)$$

Eq. (9.30a), (9.30b), and (9.31) are then six homogeneous equations in the five wave amplitudes  $\psi_i, \psi_r, \psi_t, \psi_a, \psi_b$ , from which we can extract the two desired *reciprocity relations*:

$$\boxed{\mathbf{r}' = -\mathbf{r}, \quad \mathbf{t}\mathbf{t}' - \mathbf{r}\mathbf{r}' = 1}. \quad (9.32)$$

Since there is no mechanism to produce a phase shift as the waves propagate across a perfectly sharp boundary, it is reasonable to expect  $\mathbf{r}$ ,  $\mathbf{r}'$ ,  $\mathbf{t}$  and  $\mathbf{t}'$  all to be real, as indeed they are (Ex. 9.10). (If the interface has a finite thickness, it is possible to adjust the spatial origins on the two sides of the interface so as to make  $\mathbf{r}$ ,  $\mathbf{r}'$ ,  $\mathbf{t}$  and  $\mathbf{t}'$  all be real, leading to the reciprocity relations (9.32), but a price will be paid; see Ex. 9.10.)

Return, now, to the case of finite slab thickness. By solving Eqs. (9.30) for the reflected and transmitted fields and invoking the reciprocity relations (9.32), we obtain

$$\frac{\psi_r}{\psi_i} \equiv \mathbf{r}_e = \frac{\mathbf{r}(1 - e^{i\varphi})}{1 - \mathbf{r}^2 e^{i\varphi}}, \quad \frac{\psi_t}{\psi_i} \equiv \mathbf{t}_e = \frac{(1 - \mathbf{r}^2)e^{i\varphi/(2\cos^2\theta)}}{1 - \mathbf{r}^2 e^{i\varphi}}. \quad (9.33a)$$

Here  $\mathbf{r}_e$  and  $\mathbf{t}_e$  are *the etalon's reflection and transmission coefficients*; and  $\varphi = k(2s_1 - s_2)$ , which reduces to

$$\varphi = 2n\omega d \cos\theta/c, \quad (9.33b)$$

is the light's *round-trip phase shift* (along path  $a$  then  $b$ ) *inside the etalon*, relative to the phase of the incoming light that it meets at location  $A$ . If  $\varphi$  is a multiple of  $2\pi$ , the round-trip light will superpose coherently on the new, incoming light.

We are particularly interested in the *reflectivity and transmissivity* for the energy flux, i.e. the coefficients that tell us what fraction of the total flux (and therefore also the total power) incident on the etalon is reflected by it, and what fraction emerges from its other side:

$$R = |\mathbf{r}_e|^2 = \frac{|\psi_r|^2}{|\psi_i|^2} = \frac{2\mathbf{r}^2(1 - \cos \varphi)}{1 - 2\mathbf{r}^2 \cos \varphi + \mathbf{r}^4}, \quad T = |\mathbf{t}_e|^2 = \frac{|\psi_t|^2}{|\psi_i|^2} = \frac{(1 - \mathbf{r}^2)^2}{1 - 2\mathbf{r}^2 \cos \varphi + \mathbf{r}^4}. \quad (9.33c)$$

From these expressions, we see that

$$\boxed{R + T = 1}, \quad (9.33d)$$

which says that the energy flux reflected from the slab plus that transmitted is equal to that impinging on the slab (energy conservation). It is actually the reciprocity relations (9.32) for the amplitude reflection and transmission coefficients that have enforced this energy conservation. If they had contained a provision for absorption or scattering of light in the interfaces,  $R + T$  would have been less than one.

We shall discuss these reflection and transmission coefficients at length in the next subsection; but first, in a set of example exercises, we shall clarify some important issues related to the above analysis.

\*\*\*\*\*

## EXERCISES

### Exercise 9.10 *Example: Reciprocity Relations for a Locally Planar Optical Device*

Modern mirrors, etalons, beam splitters, and other optical devices are generally made of glass or fused silica (quartz), with dielectric coatings on their surfaces. The coatings consist of alternating layers of materials with different dielectric constants, so the index of refraction  $n$  varies periodically. If, for example, the period of  $n$ 's variations is half a wavelength of the radiation, then waves reflected from successive dielectric layers build up coherently, producing a large net reflection coefficient; the result is a highly reflecting mirror.

In this exercise, we shall use a method due to Stokes to derive the reciprocity relations for devices with dielectric coatings, and in fact for much more general devices. Specifically, our derivation will be valid for *locally plane-fronted, monochromatic waves impinging on an arbitrary, locally planar, lossless optical device*. [By "locally" plane-fronted and planar, we mean that transverse variations are on scales sufficiently long compared to the wavelength of light that we can use the plane-wave analysis sketched below; for example, the spherical mirrors and Gaussian beams of an interferometric gravitational-wave detector (Fig. 9.13) easily satisfy this requirement. By lossless we mean that there is no absorption or scattering of the light.] The device could be a mirror, a surface with an antireflection coating (Ex. 9.13 below), an etalon, or any sequence of such objects with locally parallel surfaces.

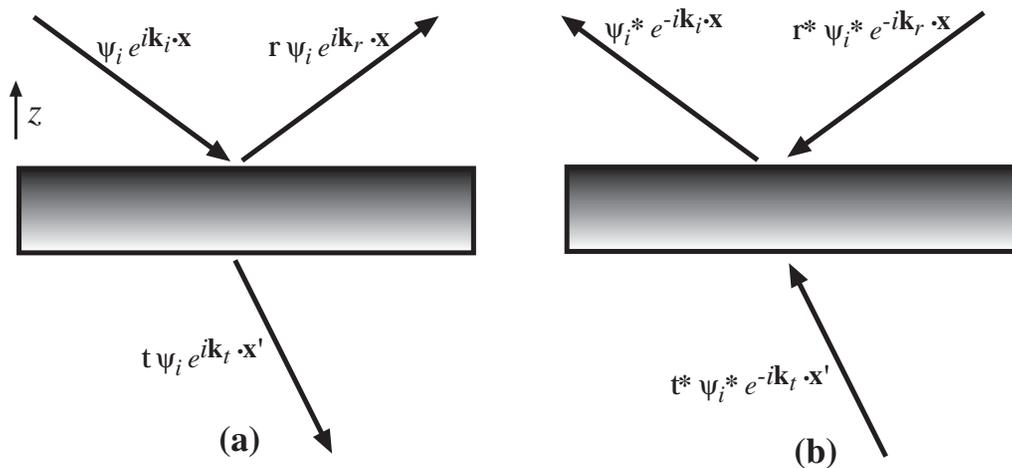
Let a plane, monochromatic wave  $\psi_i e^{i\mathbf{k}_i \cdot \mathbf{x}} e^{-i\omega t}$  impinge on the optical device from above, and orient the device so its normal is in the  $z$  direction and it is translation invariant in the  $x$  and  $y$  directions; see Fig. 9.8a. Then the reflected and transmitted waves are as shown in the figure. Because the medium below the device can have a different index of refraction from that above, the waves' propagation direction below may be different from that above, as shown. For reasons explained in part (e) below, we denote position below the device by  $\mathbf{x}'$  and position above the device by  $\mathbf{x}$ . Some arbitrary choice has been made for the locations of the vertical origins  $z = 0$  and  $z' = 0$  on the two sides of the device.

- (a) Consider a thought experiment in which the waves of Fig. 9.8a are time-reversed, so they impinge on the device from the original reflection and transmission directions and emerge toward the original input direction, as shown in Fig. 9.8b. If the device had been lossy, the time-reversed waves would not satisfy the field's wave equation; the absence of losses guarantees they do. Show that, mathematically, the time reversal can be achieved by complex conjugating the spatial part of the waves, while leaving the temporal part  $e^{-i\omega t}$  unchanged. (Such phase conjugation can be achieved in practice using techniques of nonlinear optics, as we shall see in the next chapter.) Show, correspondingly, that the spatial part of the time-reversed waves is described by the formulas shown in Fig. 9.8b.
- (b) Use the reflection and transmission coefficients to compute the waves produced by the inputs of Fig. 9.8b. From the requirement that the wave emerging from the device's upward side have the form shown in the figure, conclude that

$$\boxed{1 = r r^* + t' t'^*} . \quad (9.34a)$$

Similarly, from the requirement that no wave emerge from the device's downward side, conclude that

$$\boxed{0 = t t^* + t' r'^*} . \quad (9.34b)$$



**Fig. 9.8:** Construction for deriving reciprocity relations for amplitude transmission and reflection coefficients.

Eqs. (9.34) are the most general form of the reciprocity relations for lossless, planar devices.

- (c) For a sharp interface between two homogeneous media, combine these general reciprocity relations with the ones derived in the text, Eq. (9.32), to show that  $\mathbf{t}$ ,  $\mathbf{t}'$ ,  $\mathbf{r}$  and  $\mathbf{r}'$  are all real (as was asserted in the text).
- (d) For the etalon of Figs. 9.6 and 9.7,  $\mathbf{r}_e$  and  $\mathbf{t}_e$  are given by Eqs. (9.33a). What, do the reciprocity relations tell us about the coefficients for light propagating in the opposite direction,  $\mathbf{r}'_e$  and  $\mathbf{t}'_e$ ?
- (e) Show that for a general optical device, the reflection and transmission coefficients can all be made real by appropriate, independent adjustments of the origins of the vertical coordinates  $z$  (for points above the device) and  $z'$  (for points below the device). More specifically, show that by setting  $z_{\text{new}} = z_{\text{old}} + \delta z$  and  $z'_{\text{new}} = z'_{\text{old}} + \delta z'$  and choosing  $\delta z$  and  $\delta z'$  appropriately, one can make  $\mathbf{t}$  and  $\mathbf{r}$  real. Show further that the reciprocity relations (9.34a), (9.34b) then imply that  $\mathbf{t}'$  and  $\mathbf{r}'$  are also real. Finally, show that *this adjustment of origins brings the real reciprocity relations into the same form (9.32) as for a sharp interface between two homogeneous media.*

As attractive as it may be to have these coefficients real, one must keep in mind some disadvantages: (i) the displaced origins for  $z$  and  $z'$  in general will depend on frequency, and correspondingly (ii) frequency-dependent information (most importantly, *frequency-dependent phase shifts of the light*) are lost by making the coefficients real. If the phase shifts depend only weakly on frequency over the band of interest (as is typically the case for the dielectric coating of a mirror face), then these disadvantages are unimportant and it is conventional to choose the coefficients real. If the phase shifts depend strongly on frequency over the band of interest [e.g., for the etalon of Eqs. (9.33a), when its two faces are highly reflecting and its round-trip phase  $\varphi$  is near a multiple of  $2\pi$ ], the disadvantages are severe. One then should leave the origins frequency independent, and correspondingly leave the device's  $\mathbf{r}$ ,  $\mathbf{r}'$ ,  $\mathbf{t}$  and  $\mathbf{t}'$  complex [as we have for the etalon, in Eqs. (9.33a)].

**Exercise 9.11** *\*\*Example: Transmission and Reflection Coefficients for an Interface Between Dielectric Media*

Consider monochromatic electromagnetic waves that propagate from a medium with index of refraction  $\mathbf{n}_1$  into a medium with index of refraction  $\mathbf{n}_2$ . Let  $z$  be a cartesian coordinate perpendicular to the planar interface between the medium.

- (a) From the Helmholtz equation  $[-\omega^2 + (c^2/\mathbf{n}^2)\nabla^2]\psi = 0$ , show that both  $\psi$  and  $\psi_{,z}$  must be continuous across the interface.
- (b) Using these continuity requirements, show that for light that propagates orthogonal to the interface ( $z$  direction), the reflection and transmission coefficients, in going from medium 1 to medium 2, are

$$\boxed{\mathbf{r} = \frac{\mathbf{n}_1 - \mathbf{n}_2}{\mathbf{n}_1 + \mathbf{n}_2}, \quad \mathbf{t} = \frac{2\mathbf{n}_1}{\mathbf{n}_1 + \mathbf{n}_2}}. \quad (9.35)$$

Notice that these  $\mathbf{r}$  and  $\mathbf{t}$  are both real.

- (c) Use the reciprocity relations (9.34) to deduce the reflection and transmission coefficients  $\mathbf{r}'$  and  $\mathbf{t}'$  for a wave propagating in the opposite direction, from medium 2 to medium 1.

**Exercise 9.12** \*\*\* *Example: Etalon's Light Fields Computed by Summing the Contributions from a Sequence of Round Trips*

Study the step-by-step build up of the field inside an etalon and the etalon's transmitted field, when the input field is suddenly turned on. More specifically:

- (a) When the wave first turns on, the transmitted field inside the etalon, at point  $A$  of Fig. 9.7, is  $\psi_a = \mathbf{t}\psi_i$ , which is very small if the reflectivity is high so  $|\mathbf{t}| \ll 1$ . Show (with the aid of Fig. 9.7) that, after one round-trip-travel time in the etalon, the transmitted field at  $A$  is  $\psi_a = \mathbf{t}\psi_i + (\mathbf{r}')^2 e^{i\varphi} \mathbf{t}\psi_i$ . Show that for high reflectivity and on resonance, the tiny transmitted field has doubled in amplitude and its energy flux has quadrupled.
- (b) Compute the transmitted field  $\psi_a$  at  $A$  after more and more round trips, and watch it build up. Sum the series to obtain the steady-state field  $\psi_a$ . Explain the final, steady state amplitude: why is it not infinite, and why, physically, does it have the value you have derived.
- (c) Show that, at any time during this buildup, the field transmitted out the far side of the etalon is  $\psi_t = \mathbf{t}'\psi_a e^{iks_1}$  [Eq. (9.30b)]. What is the final, steady-state transmitted field? Your answer should be Eq. (9.33a).

**Exercise 9.13** \*\*\* *Example: Anti-reflection Coating*

A common technique used to reduce the reflection at the surface of a lens is to coat it with a quarter wavelength of material with refractive index equal to the geometric mean of the refractive indices of air and glass.

- (a) Show that this does indeed lead to perfect transmission of normally incident light.
- (b) Roughly how thick must the layer be to avoid reflection of blue light? Estimate the energy-flux reflection coefficient for red light in this case.

Note: The amplitude reflection coefficient at an interface is given by Eq. (9.35).

**Exercise 9.14** *Problem: Oil Slick*

When a thin layer of oil lies on top of water, one sometimes sees beautiful, multicolored, irregular bands of light reflecting off the oil layer. Explain what causes this.

\*\*\*\*\*

### 9.4.2 Fabry-Perot Interferometer, and Box on Modes of a Fabry-Perot Cavity with Spherical Mirrors

When an etalon's two faces are highly reflecting (reflection coefficient  $\mathfrak{r}$  near unity), we can think of them as *mirrors*, between which the light resonates. The etalon is then a special case of a *Fabry-Perot interferometer*. The general case is any device in which light resonates between two high-reflectivity mirrors. The mirrors need not be planar and need not have the same reflectivities, and the resonating light need not be plane fronted.

A common example is the *optical cavity* of Fig. 7.9, formed by two mirrors that are segments of spheres, which we studied using geometric optics in Ex. 7.12. Because the phase fronts of a Gaussian beam (Sec. 8.5.5) are also spherical, such a beam can resonate in the optical cavity, if (i) the beam's waist location and waist radius are adjusted so its phase-front radii of curvature, at the mirrors, are the same as the mirrors' radii of curvature, and (ii) the light's frequency is adjusted so a half integral number of wavelengths fit perfectly inside the cavity. Box 9.3 gives details for the case where the two mirrors have identical radii of curvature. In that box we also learn that the Gaussian beams are not the only eigenmodes that can resonate inside such a cavity. Other, "higher-order" modes can also resonate. They have more complex transverse distributions of the light. There are two families, of such modes: one with rectangular transverse light distributions, and the other with wedge-shaped, spoke-like light distributions.

For any Fabry-Perot interferometer with identical mirrors, driven by light with a transverse cross section that matches one of the interferometer's modes, one can study the interferometer's response to the driving light by the same kind of analysis as we used for an etalon in the previous section; and the result will be the same: The interferometer's reflected and transmitted light, at a given transverse location,  $\{x, y\}$ , will be given by

$$\boxed{\frac{\psi_r}{\psi_i} \equiv \mathfrak{r}_{\text{FP}} = \frac{\mathfrak{r}(1 - e^{i\varphi})}{1 - \mathfrak{r}^2 e^{i\varphi}}, \quad \frac{\psi_t}{\psi_i} \equiv \mathfrak{t}_{\text{FP}} = \frac{(1 - \mathfrak{r}^2)e^{i\varphi/2}}{1 - \mathfrak{r}^2 e^{i\varphi}}}. \quad (9.36)$$

[Eqs. (9.33a) with  $\theta = 0$  so the light rays are orthogonal to the mirrors]. Here  $\mathfrak{r}$  is the mirrors' reflection coefficient, and the round-trip phase is now

$$\boxed{\varphi = 2\pi\omega/\omega_f + \varphi_G, \quad \text{where} \quad \omega_f = 2\pi/\tau_{\text{rt}}}. \quad (9.37)$$

Here  $\tau_{\text{rt}}$  is the time required for a high-frequency photon to travel round-tip in the interferometer, along the optic axis, from one mirror to the other;  $\omega_f$  (called the *free spectral range*) is, as we shall see, the angular-frequency separation between the interferometer's resonances; and  $\varphi_G$  is an additive contribution (called the *Gouy phase*), caused by the curvature of the phase fronts [e.g. the  $\tan^{-1}(z/z_o)$  term in Eq. (8.40a) for a Gaussian beam and in Eqs. (2) and (1) of Box 9.3 for higher-order modes]. Because  $\varphi_G$  is of order one while  $2\pi\omega/\omega_{f_{sr}}$  is huge compared to one, and because  $\varphi_G$  changes very slowly with changing light frequency, it is unimportant in principle (and we henceforth shall ignore it). However, it is important in practice: it causes modes with different transverse light distributions (e.g. the Gaussian and higher order modes in Box 9.3), which have different Gouy phases, to resonate at different frequencies.

### Box 9.3

#### Modes of a Fabry-Perot Cavity with Spherical Mirrors

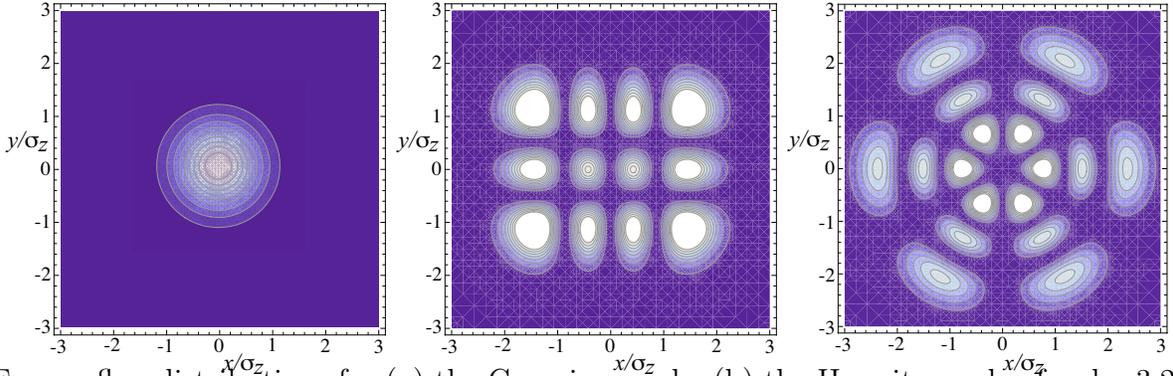
Consider a Fabry-Perot cavity whose spherical mirrors have the same radius of curvature  $R$  and are separated by a distance  $L$ . Introduce (i) Cartesian coordinates with  $z = 0$  at the cavity's center, and (ii) the same functions as we used for Gaussian beams [Eqs. (8.40b)]:

$$z_0 = \frac{k\sigma_0^2}{2} = \frac{\pi\sigma_0^2}{\lambda}, \quad \sigma_z = \sigma_0(1 + z^2/z_0^2)^{1/2}, \quad R_z = z(1 + z_0^2/z^2), \quad (1)$$

with  $k$  the wave number and  $\sigma_0$  a measure of the transverse size of the beam at the cavity's center. Then it is straightforward to verify that the following functions (i) satisfy the Helmholtz equation, (ii) are orthonormal when integrated over their transverse Cartesian coordinates  $x$  and  $y$ , and (iii) have phase fronts (surfaces of constant phase) that are spheres with radius of curvature  $R_z$ :

$$u_{nm}(x, y, z) = \frac{e^{-(x^2+y^2)/\sigma_z^2}}{\sqrt{2^{m+n-1}\pi m!n!} \sigma_z} H_m\left(\frac{\sqrt{2}x}{\sigma_z}\right) H_n\left(\frac{\sqrt{2}y}{\sigma_z}\right) \times \exp\left\{i\left[\frac{k(x^2+y^2)}{2R_z} + kz - (n+m+1)\tan^{-1}\frac{z}{z_0}\right]\right\}. \quad (2)$$

Here  $H_n(\xi) = e^{\xi^2} d^n e^{-\xi^2} / d\xi^n$  is the Hermite polynomial and  $m$  and  $n$  range over non-negative integers. By adjusting  $\sigma_0$ , we can make the phase-front radius of curvature  $R_z$  match that,  $R$ , of the mirrors at the mirror locations,  $z = \pm L/2$ . Then  $u_{mn}$  are a transversely orthonormal set of modes for the light field inside the cavity. Their flux distribution  $|u_{mn}|^2$  on each mirror consists of an  $m+1$  by  $n+1$  matrix of discrete spots; see drawing (b) below. The mode with  $m = n = 0$  [drawing (a)] is the Gaussian beam explored in the previous chapter: Eqs. (8.40). A given mode, specified by  $\{m, n, k\}$  cannot resonate inside the cavity unless its wave number matches the cavity length, in the sense that the the total phase shift in traveling from the cavity center  $z = 0$  to the cavity end, at  $z = \pm L/2$ , is an integral multiple of  $\pi/2$ ; i.e.,  $kL/2 - (n+m+1)\tan^{-1}(L/2z_0) = N\pi/2$ .



Energy flux distributions for (a) the Gaussian mode, (b) the Hermite mode of order 3,2, (c) the Laguerre mode of order 2,3. The contours are at 90%, 80%, ..., 10% of maximum.

**Box 9.3 (Continued)**

There is a second family of modes that can resonate in the cavity, one whose eigenfunctions separate in circular polar coordinates:

$$u_{pm}(\varpi, \phi, z) = \frac{2p! e^{-\varpi^2/\sigma_z^2}}{\sqrt{1 + \delta_{m0}} \pi (p+m)! \sigma_z} \left( \frac{\sqrt{2}\varpi}{\sigma_z} \right)^m L_p^m \left( \frac{2\varpi^2}{\sigma_z^2} \right) \begin{pmatrix} \cos m\phi \text{ or} \\ \sin m\phi \end{pmatrix} \\ \times \exp \left\{ i \left[ \frac{k\varpi^2}{2R_z} + kz - (2p+m+1) \tan^{-1} \frac{z}{z_0} \right] \right\}, \quad (1)$$

where  $L_p^m(\xi)$  is the associated Laguerre polynomial and  $p$  and  $m$  range over non-negative integers. These modes make spots on the mirrors shaped like azimuthal wedges, cut radially by circles; see drawing (c) above. Again, they can resonate only if the phase change from the center of the cavity to an end mirror at  $z = \pm L/2$  is an integral multiple of  $\pi/2$ ; i.e.,  $kL/2 - (2p+m+1) \tan^{-1}(L/2z_0) = N\pi/2$ .

As one goes to larger mode numbers  $m, n$  (Hermite modes) or  $p, m$  (Laguerre modes), the region with substantial light power gets larger; see drawings above — and, as a result, more light gets lost off the edges of the cavity's mirrors. As a result, unless the mirrors are made very large, high-order modes have large losses and do not resonate well.

For further detail on these modes, see, e.g., Secs. 2.5–2.8 and 4.3 of Yariv and Yeh (2007).

The Fabry-Perot interferometer's power transmissivity  $T$  and reflectivity  $R$  are given by Eq. (9.33c), which we can rewrite in the following, simpler form:

$$T = 1 - R = \frac{1}{1 + (2\mathcal{F}/\pi)^2 \sin^2 \frac{1}{2}\varphi}. \quad (9.38)$$

Here  $\mathcal{F}$ , called the interferometer's *finesse*, is defined by

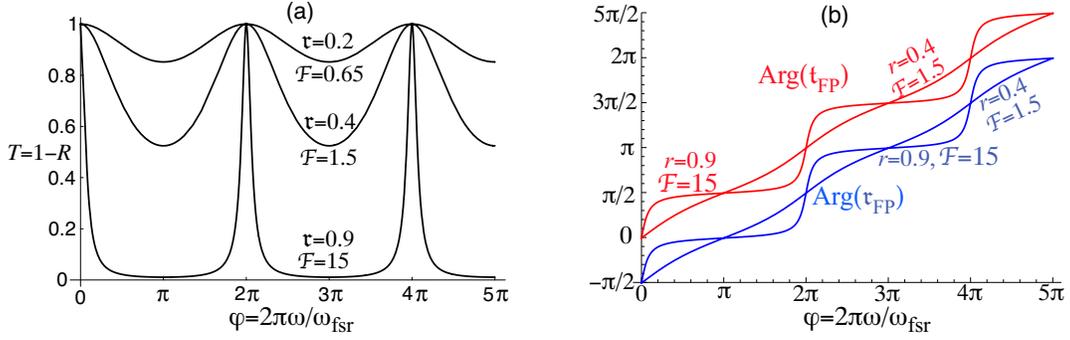
$$\mathcal{F} \equiv \pi \mathfrak{r} / (1 - \mathfrak{r}^2), \quad (9.39)$$

[This *finesse* should not be confused with the *coefficient of finesse*  $F = (2\mathcal{F}/\pi)^2$ , which is sometimes used in optics but which we shall eschew so as to avoid confusion.]

In Fig. 9.9 we plot, as functions of the round-trip phase  $\varphi = 2\pi\omega/\omega_f$  (ignoring  $\varphi_G$ ), the interferometer's power reflectivity and transmissivity  $T$  and  $R$ , and the phase changes  $\arg(\mathfrak{t}_{FP})$  and  $\arg(\mathfrak{r}_{FP})$  of the light that is transmitted and reflected by the interferometer.

Notice in Fig. 9.9a that, when the *finesse*  $\mathcal{F}$  is large compared to unity, the interferometer exhibits sharp resonances at frequencies separated by the free spectral range  $\omega_{fsr}$ . On resonance, the interferometer is perfectly transmitting,  $T = 1$ ; away from resonance, it is nearly perfectly reflecting,  $R \simeq 1$ . The *full width at half maximum* (“half width”) of each sharp transmission resonance is given by

$$\delta\varphi_{1/2} = \frac{2\pi}{\mathcal{F}}, \quad \delta\omega_{1/2} = \frac{\omega_f}{\mathcal{F}}. \quad (9.40a)$$



**Fig. 9.9:** (a) Power transmissivity and reflectivity [Eq. (9.38)] for a Fabry-Perot interferometer with identical mirrors that have reflection coefficients  $\tau$ , as a function of the round-trip phase shift  $\varphi$  inside the interferometer. (b) The phase of the light transmitted (red) or reflected (blue) from the interferometer, relative to the input phase [Eqs. (9.36)]. The interferometer’s finesse  $\mathcal{F}$  is related to the mirrors’ reflectivity by Eq. (9.39).

In other words, if the frequency  $\omega$  of the light is swept slowly through resonance, the transmission will be within 50% of its peak value (unity) over a bandwidth  $\delta\omega_{1/2} = \omega_f/\mathcal{F}$ . Notice also, in Fig. 9.9b, that for large finesse, near resonance the phase of the reflected and transmitted light changes very rapidly with a change in frequency of the driving light. Precisely on resonance, that rate of change is

$$\left(\frac{d \arg(t_{\text{FP}})}{d\omega}\right)_{\text{on resonance}} = \left(\frac{d \arg(r_{\text{FP}})}{d\omega}\right)_{\text{on resonance}} = \frac{2\mathcal{F}}{\omega_f} = \frac{2}{\delta\omega_{1/2}}. \quad (9.40b)$$

The large transmissivity at resonance, for large finesse, can be understood by considering what happens when one first turns on the incident wave. Since the reflectivity of the first (input) mirror is near unity, the incoming wave has a large amplitude for reflection, and correspondingly only a tiny amplitude for transmission into the optical cavity. The tiny bit that gets transmitted travels through the first mirror, gets strongly reflected from the second mirror, and returns to the first precisely in phase with the incoming wave (because  $\varphi$  is an integer multiple of  $2\pi$ ). Correspondingly, it superposes coherently on the tiny field being transmitted by the incoming wave, and so the net wave inside the cavity is doubled. After one more round trip inside the slab, this wave returns to the first face again in phase with the tiny field being transmitted by the incoming wave; again they superpose coherently; and the internal wave now has a three times larger amplitude than it began with. This process continues until a very strong field has built up inside the cavity (Ex. 9.12). As it builds up, that field begins to leak out of the cavity’s first mirror with just such a phase as to destructively interfere with the wave being reflected there. The net reflected wave is thereby driven close to zero. The field leaking out of the second mirror has no other wave to interfere with. It remains strong, so the interferometer settles down into a steady state with strong net transmission. Heuristically, one can say that, because the wave inside the cavity is continually constructively superposing on itself, the cavity “sucks” almost all the incoming wave into itself, and then ejects it out the other side. (Quantum mechanically, this sucking is due to the photons’ Bose-Einstein statistics: the photons “want” to be in the same quantum

state. We shall study this phenomenon in the context of plasmons that obey Bose-Einstein statistics, in Chap. 23 [paragraph containing Eq. (23.46)].

This discussion makes clear that, when the properties of the input light are changed, a high-finesse Fabry-Perot interferometer will change its response rather slowly — on a timescale approximately equal to the inverse of the resonance half width, i.e. the finesse times the round-trip travel time:

$$\tau_{\text{response}} \sim \frac{1}{\delta\omega_{1/2}} = \mathcal{F} \tau_{\text{rt}}. \quad (9.40c)$$

These properties of a high-finesse Fabry-Perot interferometer are very similar to those of a high- $Q$  mechanical or electrical oscillator. The similarity arises because, in both cases, energy is being stored in a resonant, sinusoidal manner inside the device (the oscillator or the interferometer). For the interferometer, the light's round-trip travel time,  $\tau_{\text{rt}}$ , is analogous to the oscillator's period, the interferometer's free spectral range,  $\omega_f$ , is analogous to the oscillator's resonant angular frequency, and the interferometer's finesse,  $\mathcal{F}$ , is analogous to the oscillator's quality factor,  $Q$ . However, there are some major differences between an ordinary oscillator and a Fabry-Perot interferometer. Perhaps the most important is that the interferometer has several large families of resonant modes (families characterized by the number of longitudinal nodes between the mirrors, and by the two-dimensional transverse distributions of the light), whereas an oscillator has just one mode. This gives an interferometer much greater versatility than a simple oscillator.

### 9.4.3 Fabry-Perot Applications: spectrometer, laser, mode-cleaning cavity, beam-shaping cavity, PDH laser stabilization, optical frequency comb

Just as mechanical and electrical oscillators have a wide variety of important applications in science and technology, so also do Fabry-Perot interferometers, but more so. In this section, we shall sketch a few of them.

#### *Spectrometer*

In the case of a Fabry-Perot etalon (highly reflecting parallel mirrors; plane-parallel light beam), the resonant transmission enables the etalon to be used as a spectrometer. The round-trip phase change  $\varphi = 2n\omega d \cos \theta / c$  inside the etalon varies linearly with the wave's angular frequency  $\omega$ , but only waves with round-trip phase  $\varphi$  near an integer multiple of  $2\pi$  will be transmitted efficiently. The etalon can be tuned to a particular frequency by varying either the slab width  $d$  or the angle of incidence of the radiation (and thence the angle  $\theta$  inside the etalon). Either way, impressively good chromatic resolving power can be achieved. We say that waves with two nearby frequencies can just be resolved by an etalon when the half power point of the transmission coefficient of one wave coincides with the half power point of the transmission coefficient of the other. Using Eq. (9.38) we find that the phases for the two frequencies must differ by  $\delta\varphi \simeq 2\pi/\mathcal{F}$ ; and correspondingly, since  $\varphi = 2n\omega d \cos \theta / c$ , the *chromatic resolving power* is

$$\mathcal{R} = \frac{\omega}{\delta\omega} = \frac{4\pi n d \cos \theta}{\lambda_{\text{vac}} \delta\varphi} = \frac{2n d \cos \theta \mathcal{F}}{\lambda_{\text{vac}}}. \quad (9.41)$$

Here  $\lambda_{\text{vac}} = 2\pi c/\omega$  is the wavelength in vacuum — i.e. outside the etalon.

### ***Laser***

Fabry-Perot interferometers are exploited in the construction of many types of lasers. For example, in a gas phase laser, the atoms are excited to emit a spectral line. This radiation is spontaneously emitted isotropically over a wide range of frequencies. Placing the gas between the mirrors of a Fabry-Perot interferometer allows one or more highly collimated and narrow-band modes to be trapped and, while trapped, to be amplified by stimulated emission, i.e. to lase. See Sec. 10.2.1 of the next chapter.

### ***Mode Cleaner for a Messy Laser Beam***

The output beam from a laser often has a rather messy cross-sectional profile, e.g. due to containing a number of modes of excitation of the Fabry-Perot interferometer in which the lasing material is contained (cf. the discussion of possible modes in Box 9.3). For many applications, one needs a much cleaner laser beam, e.g., one with a Gaussian profile. To clean the beam, one can send it into a high-finesse Fabry-Perot cavity with identical spherical mirrors, whose mirror curvatures and cavity length are adjusted so that, among the modes present in the beam, only the desired Gaussian mode will resonate and thereby be transmitted (see the sharp transmission peaks in Fig. 9.9a above). The beam's unwanted other modes will not resonate in the cavity, and therefore will be reflected backward off its input mirror, leaving the transmitted beam clean.

### ***Beam Shaping Cavity***

In some applications one wants a light beam whose cross sectional distribution of flux  $F(x, y)$  is different from any of the modes that resonate in a spherical-mirror cavity — for example, one might want a circular, flat-topped flux distribution  $F(\varpi)$  with steeply dropping edges, like the shape of a circular mesa in a North American desert. One can achieve the desired light distribution, or something approximating it, as follows: Build a Fabry-Perot cavity with identical mirrors that are shaped in such a way that there is a cavity mode with the desired distribution. Then drive the cavity with a Gaussian beam. That portion of the beam which has the desired flux distribution will resonate in the interferometer and leak out of the other mirror as the desired beam; the rest of the input beam will be rejected by the cavity.

### ***Laser Stabilization***

There are two main ways to stabilize the frequency of a laser. One is to lock it onto the frequency of some fundamental atomic or molecular transition. The other is to lock it onto a resonant frequency of a mechanically stable Fabry-Perot cavity—a technique called Pound-Drever-Hall (PDH) locking.

In PDH locking to a cavity with identical mirrors, one passes the laser's output light (with frequency  $\omega$ ) through a device that modulates its frequency,<sup>6</sup> so  $\omega$  becomes  $\omega + \delta\omega$  with  $\delta\omega = \sigma \cos(\Omega t)$  and  $\sigma \ll \delta\omega_{1/2}$ , the cavity resonance's half width. One then sends the modulated light into the cavity and monitors the reflected light power. Assume, for simplicity, that the modulation is slow compared to the cavity response time,  $\Omega \ll 1/\tau_{\text{response}}$ .

---

<sup>6</sup>Actually, one sends it through a phase modulator called a *Pockels cell*, consisting of a crystal whose index of refraction is modulated by applying an oscillating voltage to it. The resulting phase modulation  $\delta\phi \propto \sin\omega t$  is equivalent to a frequency modulation  $\delta\omega = d\delta\phi/dt \propto \cos\Omega t$ .

Then the cavity's response at any moment will be that for steady light, i.e. the reflected power will be  $P_i R(\omega + \delta\omega)$ , where  $R$  is the reflectivity at frequency  $\omega + \delta\omega$ . Using Eq. (9.38) for the reflectivity, specialized to frequencies very near resonance so the denominator is close to one, and using Eqs. (9.37) and (9.40a), we bring this into the form

$$P_r = P_i R(\omega + \delta\omega) = P_i \times \left[ R(\omega) + \frac{dR}{d\omega} \delta\omega(t) \right] = P_i \left[ R(\omega) + \frac{8\sigma(\omega - \omega_o)}{(\delta\omega_{1/2})^2} \cos \Omega t \right], \quad (9.42)$$

where  $\omega_o$  is the cavity's resonant frequency (at which  $\varphi$  is an integral multiple of  $2\pi$ ).

The modulated part of the reflected power has an amplitude directly proportional to the laser's frequency error,  $\omega - \omega_o$ . In the PDH technique, one monitors this modulated power with a photodetector, followed by a band-pass filter on the photodetector's output electric current to get rid of the unmodulated part of the signal [arising from  $P_i R(\omega)$ ]. The amplitude of the resulting, modulated output current is proportional to  $\omega - \omega_o$  and is used to control a feedback circuit that drives the laser back toward the desired, cavity-resonant frequency  $\omega_o$ . See, e.g., Black (2001) for details.

In Ex. 9.15 it is shown that, if one needs a faster feedback and therefore requires a modulation frequency  $\Omega \gtrsim 1/\tau_{\text{response}}$ , this PDH locking technique still works.

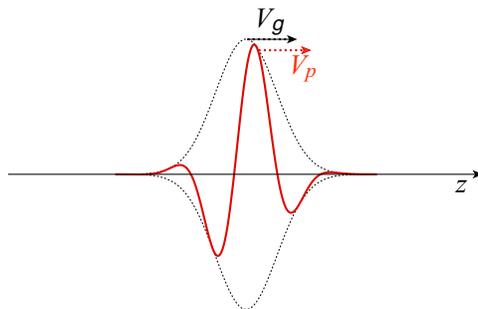
This technique was invented by Ronald Drever for use in interferometric gravitational-wave detectors, relying on earlier ideas of Robert Pound, and it was first demonstrated experimentally by Drever and John Hall. It is now used widely in many areas of science and technology.

### ***Optical Frequency Comb***

John Hall and Theodor Hänsch were awarded the 2005 Nobel Prize for development of the optical frequency comb. This powerful tool is based on an optical cavity of length  $L$ , filled with a lasing medium that creates and maintains a sharply pulsed internal light field with the following form (Fig. 9.10):

$$\psi = \psi_o(z - V_g t) \exp[ik_p(z - V_p t)]. \quad (9.43)$$

Here (i) we use a  $z$  coordinate that increases monotonically along the optic axis as one moves rightward from mirror 1 to mirror 2, then leftward from 2 to 1, then rightward from 1 to 2, etc.; (ii)  $\exp[ik_p(z - V_p t)]$  is the  $p$ 'th longitudinal monochromatic mode of the cavity, with



**Fig. 9.10:** The sharply pulsed electric field (9.43) inside a Fabry Perot cavity. The envelope is shown dotted; the red curve is the full field  $\psi$ .

wave number  $k_p \equiv p\pi/L$ , phase velocity  $V_p$ , and angular frequency  $k_p V_p$  lying in the optical range  $\sim 10^{15}$  Hz; (iii)  $\psi_o(z - V_g t)$  is the envelope of a wave packet so narrow that only  $\sim$ one wavelength of the mode  $p$  can fit inside it; (iv) the envelope travels with group velocity  $V_g$  and does not spread.

In order that the wave packet not spread,  $V_g$  must have the same, constant value over all frequencies contained in the packet, which means that the dispersion relation must have  $0 = (\partial V_g / \partial k) = \partial^2 \omega / \partial k^2$ ; i.e.,  $\omega$  must be linear in  $k$ :

$$\omega = V_g(k + \kappa) \quad (9.44)$$

for some constant  $\kappa$ , which is typically considerably smaller than the wave numbers  $k$  contained in the packet.

It was a huge technical challenge to build a lasing cavity that creates and sustains a very narrow wave packet of this sort. Two of the keys to this achievement this were (i) using a *nonlinear* lasing medium which amplifies light more strongly at high energy fluxes  $|\psi|^2$  than low and thereby tries to produce intense, short pulses rather than long, monochromatic waves, and (ii) using some trickery to assure that the lasing medium and anything else in the cavity jointly give rise to the linear dispersion relation (9.44) over a sufficiently wide frequency band. For some of the details, see Sec. 10.2.3. Because the sharp wavepacket (9.43) has fixed relationships between the phases of the various monochromatic modes that make it up, the lasing optical cavity that creates it is called a *mode-locked laser*.

As the internal field's pulse hits mirror 2 time and again, it transmits through the mirror a sequence of outgoing pulses separated by the wave packet's round-trip travel time in the cavity,  $\tau_{rt} = 2L/V_g$ . Assuming, for pedagogical clarity, a Gaussian shape for each pulse, the oscillating internal field (9.43) produces the outgoing field  $\psi \propto \sum_n \exp[-\sigma^2(t - z/c - n\tau_{rt})^2/2] \exp[-ik_p V_p(t - z/c)]$ . Here  $1/\sigma$  is the pulse length in time, and we have assumed vacuum-light-speed propagation outside the cavity. It is helpful to rewrite the frequency  $k_p V_p$  of the oscillatory piece of this field as the sum of its nearest multiple of the cavity's free spectral range,  $\omega_f = 2\pi/\tau_{rt}$ , plus a frequency shift  $\omega_s$ :  $k_p V_p = q\omega_f + \omega_s$ . The integer  $q = (\text{largest integer contained in } k_p V_p / \omega_f)$  will typically be quite close to  $p$ , and  $\omega_s$  is guaranteed to lie in the interval  $0 \leq \omega_s < \omega_f$ . The emerging electric field is then

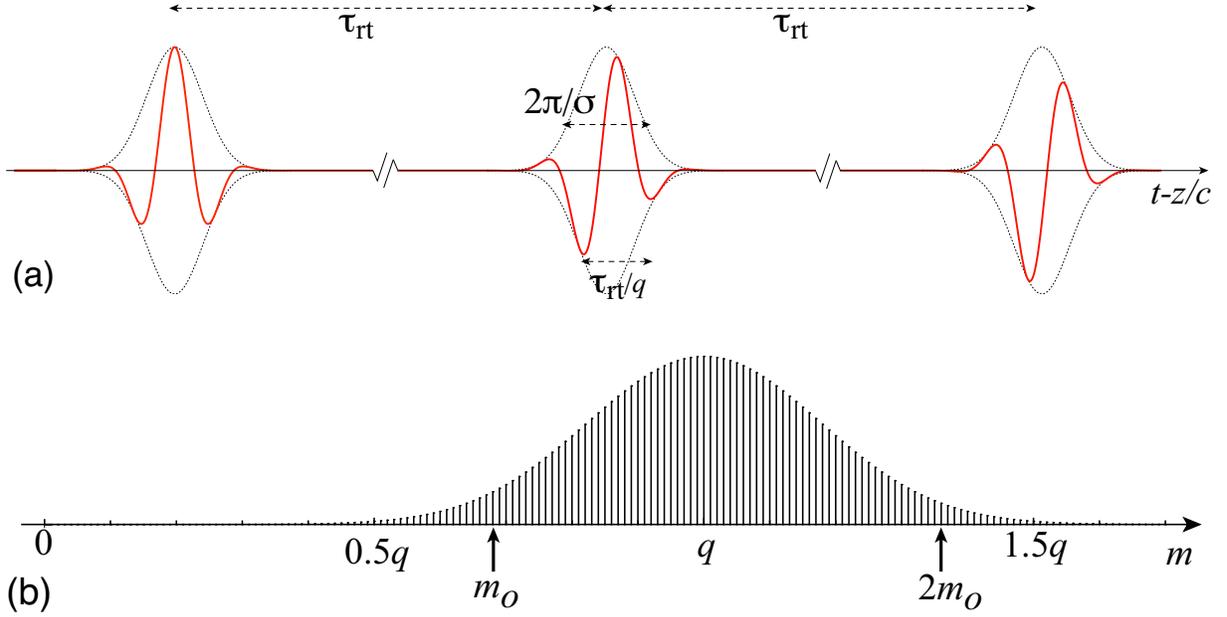
$$\psi \propto \sum_n \exp[-\sigma^2(t - z/c - n\tau_{rt})^2/2] \exp[-i(q\omega_f + \omega_s)(t - z/c)] \quad (9.45a)$$

(Fig. 9.11a).

This entire emerging field is periodic in  $t - z/c$  with period  $\tau_{rt} = 2\pi/\omega_f$ , except for the frequency-shift term  $\exp[-i\omega_s(t - z/c)]$ . The periodic piece can be expanded as a sum over discrete frequencies that are multiples of  $\omega_f = 2\pi/\tau_{rt}$ . Since the Fourier transform of a Gaussian is a Gaussian, this sum, augmented by the frequency shift term, turns out to be

$$\psi \propto \sum_{m=-\infty}^{+\infty} \exp\left[\frac{-(m - q)^2 \omega_f^2}{2\sigma^2}\right] \exp[-i(m\omega_f + \omega_s)(t - z/c)]. \quad (9.45b)$$

The set of discrete frequencies (spectral lines) appearing in this outgoing field is the *frequency comb* displayed in Fig. 9.11b.



**Fig. 9.11:** Optical frequency comb: (a) The pulsed electric field (9.45a) emerging from the cavity. (b) The field's comb spectrum; each line, labeled by  $m$ , has angular frequency  $m\omega_f$ , and is shown with height proportional to its power,  $\propto \exp[-(m-q)^2\omega_f^2/\sigma^2]$  [cf. Eq. (9.45b)].

Some concrete numbers make clear how very remarkable this pulsed electric field (9.45a) and its frequency comb (9.45b) are: (i) The Fabry-Perot cavity typically has a length  $L$  somewhere between  $\sim 3$  cm and  $\sim 3$  m, so (with the group velocity  $V_g$  of order the vacuum light speed) the round-trip travel time and free spectral range are  $\tau_{rt} = 2L/V_g \sim 0.3$  to 30 ns; and  $\omega_f/2\pi \sim 1/\tau_{rt} \sim 30$  MHz to 3 GHz, which are radio and microwave frequencies. Since the shift frequency is  $\omega_s < \omega_f$ , it is also in the radio or microwave band. (ii) The comb's central frequency is in the optical,  $q\omega_f/2\pi \sim 3 \times 10^{14}$  Hz, so the harmonic number of the central frequency is  $q \sim 10^5$  to  $10^7$ , roughly a million. (iii) The pulse width  $\sim 2/\sigma$  contains roughly one period  $2\pi/q\omega_f$  of the central frequency, so  $\sigma \sim q\omega_f/3$ , which means that most of the comb's power is contained in the range  $m \sim 2q/3$  to  $m \sim 4q/3$ ; i.e., there are roughly a million strong teeth in the comb.

It is possible to lock the comb's free spectral range  $\omega_f$  to a very good Cesium atomic clock, whose oscillation frequency  $\sim 9$ GHz is stable to  $\delta\omega/\omega \sim 10^{-12}$  (Fig. 6.11), so  $\omega_f$  has that same phenomenal stability. One can then measure the shift frequency  $\omega_s$  and calibrate the comb (identify the precise mode number  $m$  of each frequency in the comb) as follows: (i) Arbitrarily choose a tooth at the low-frequency end of the comb,  $m_o \simeq 2q/3$  (Fig. 9.11b); it has frequency  $\omega_o = m_o\omega_f + \omega_s$ . (ii) Separate the light in that tooth from light in the other teeth, and send a beam of that tooth's light through a frequency doubler (to be discussed in Sec. 10.6.1), thereby getting a beam with frequency  $2\omega_o = 2(m_o\omega_f + \omega_s)$ . (iii) By beating this beam against the light in teeth at  $m \sim 4q/3$ , identify the tooth that most closely matches this beam's frequency. It will have frequency  $2m_o\omega_f + \omega_s$ , and the frequency difference (beat frequency) will be  $\omega_s$ . This reveals  $\omega_s$  to very high accuracy; and one can count the number of teeth ( $m_o - 1$ ) between this tooth  $2m_o$  and its undoubled parent  $m_o$ , thereby learning the

precise numerical value of  $m_o$ . From this, by tooth counting, one learns the precise mode numbers  $m$  of all the optical-band teeth in the comb, and also their frequencies  $m\omega_f + \omega_s$ .

With the comb now calibrated, it can be used to measure the frequency of any other beam of light in the optical band in terms of the ticking frequency of the Cesium clock, to which the entire comb has been locked. The optical-frequency accuracies thereby achieved are orders of magnitude better than were possible before this optical frequency comb was developed. And in the near future, as optical-frequency atomic clocks become much more accurate and stable than the microwave-frequency Cesium clock (see footnote 8 in Chap. 6), this comb will be used to calibrate microwave and radio frequencies in term of the ticking rates of optical-frequency clocks.

For further details about optical frequency combs, see the review articles by Cundiff (2002) and by Cundiff and Ye (2003).

\*\*\*\*\*

## EXERCISES

### Exercise 9.15 *Problem: PDH Laser Stabilization*

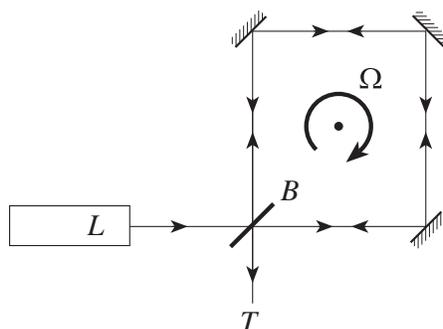
Show that the PDH method for locking a laser's frequency to an optical cavity works for modulations faster than the cavity's response time,  $\Omega \gtrsim 1/\tau_{\text{response}}$ , and even works for  $\Omega \gg 1/\tau_{\text{response}}$ . More specifically, show that the reflected power still contains the information needed for feedback to the laser. [For a quite general analysis and some experimental details, see Black (2001).]

### Exercise 9.16 *Derivation: Optical Frequency Comb*

Fill in the details of the derivation of all the equations in the section describing the optical frequency comb.

### Exercise 9.17 \*\*\* *Problem: Sagnac Interferometer*

A Sagnac interferometer is a rudimentary version of a laser gyroscope for measuring rotation with respect to an inertial frame. The optical configuration is shown in Fig. 9.12. Light



**Fig. 9.12:** Sagnac interferometer used as a type of laser gyro.

from a laser  $L$  is split by a beam splitter  $B$  and travels both clockwise and counter-clockwise around the optical circuit, reflecting off three plane mirrors. The light is then recombined at  $B$  and interference fringes are viewed through the telescope  $T$ . The whole assembly rotates with angular velocity  $\Omega$ .

Calculate the difference in the time it takes light to traverse the circuit in the two directions and show that the consequent fringe shift (total number of fringes that enter the telescope during one round trip of the light in the interferometer) can be expressed as  $\Delta N = 4A\Omega/c\lambda$ , where  $\lambda$  is the wavelength and  $A$  is the area bounded by the beams. Show further that, for a square Sagnac interferometer with side length  $L$ , the rate at which fringes enter the telescope is  $\Omega L/\lambda$ .

\*\*\*\*\*

## 9.5 T Laser Interferometer Gravitational Wave Detectors

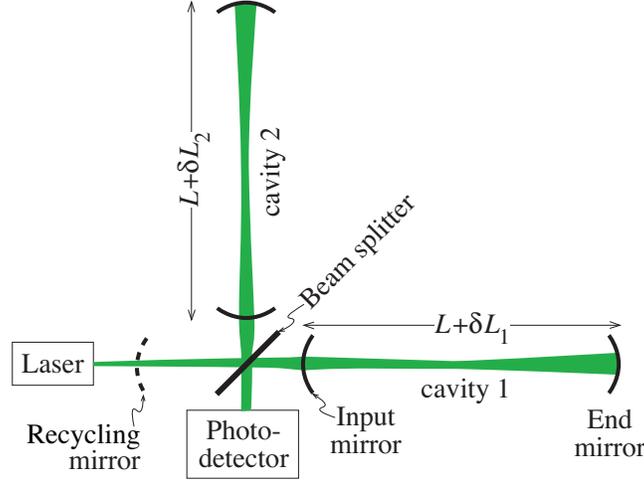
As we shall discuss in Chap. 27, gravitational waves are predicted to exist by general relativity theory, and their emission by a binary neutron-star system has already been monitored, via their back-action on the binary's orbital motion. As orbital energy is lost to gravitational waves, the binary gradually spirals inward, so its orbital angular velocity gradually increases. The measured rate of increase agrees with general relativity's predictions to within the experimental accuracy of a fraction of a percent (for which Russel Hulse and Joseph Taylor received the 1993 Nobel Prize). Unfortunately, the gravitational analog of Hertz's famous laboratory emission and detection of electromagnetic waves has not yet been performed, and cannot be in the authors' lifetime because of the waves' extreme weakness. For waves strong enough to be detectable, one must turn to violent astrophysical events, such as the collision and coalescence of two neutron stars or black holes.

When the gravitational waves reach earth and pass through a laboratory, general relativity predicts that they will produce tiny relative accelerations of free test masses. The resulting oscillatory variation of the distance between two such masses can be measured optically using a Michelson interferometer, in which (to increase the signal strength) each of the two arms is operated as a Fabry-Perot cavity. The most sensitive such gravitational wave detectors that have been operated as of 2012 are the "initial detectors" of the *Laser Interferometer Gravitational Wave Observatory* (LIGO); see LSC (2009).

In each of these gravitational-wave detectors, the two cavities are aligned along perpendicular directions as shown in Fig. 9.13. A Gaussian beam of light (Sec. 8.5.5) from a laser passes through a beam splitter, creating two beams with correlated phases. The beams excite the two cavities near resonance. Each cavity has an end mirror with extremely high reflectivity,<sup>7</sup>  $1 - \mathfrak{r}_e^2 < 10^{-4}$ , and a corner mirror ("input mirror") with a lower reflectivity,  $1 - \mathfrak{r}_i^2 \sim 0.03$ . Because of this lower reflectivity, by contrast with the etalons discussed above,

---

<sup>7</sup>Because LIGO operates with monochromatic light, it is convenient to adjust the phases of the mirrors' reflection and transmission coefficients so  $\mathfrak{r}$  and  $\mathfrak{t}$  are both real; cf. Ex. 9.11e. We do so.



**Fig. 9.13:** Schematic design of an initial gravitational wave interferometer operated in LIGO (at Livingston Louisiana and Hanford Washington, USA) during 2005–2010.

the resonant light leaks out through the input mirror instead of through the end mirror. The reflectivity of the input mirror is so adjusted that the typical photon is stored in the cavity for roughly half the period of the expected gravitational waves (a few milliseconds), which means that the input mirror's reflectivity  $\tau_i^2$ , the arm length  $L$ , and the gravitational-wave angular frequency  $\omega_{\text{gw}}$  are related by

$$\frac{L}{c(1 - \tau_i^2)} \sim \frac{1}{\omega_{\text{gw}}} . \quad (9.46)$$

The light emerging from the cavity, like that transmitted by an etalon, has a phase that is highly sensitive to the separation between the mirrors: a tiny change  $\delta L$  in their separation produces a change in the outgoing phase

$$\delta\varphi_o \simeq \frac{8\omega\delta L}{c} \frac{1}{(1 - \tau_i^2)} \sim \frac{\omega}{\omega_{\text{gw}}} \frac{\delta L}{L} \quad (9.47)$$

in the limit  $1 - \tau_i \ll 1$ ; see Ex. 9.18. The outgoing light beams from the two cavities return to the beam splitter and there are recombined. The relative distances from the beam splitter to the cavities are adjusted so that, in the absence of any perturbations of the cavity lengths, almost all the interfered light goes back toward the laser, and only a tiny (but nonzero) amount goes toward the photodetector of Fig. 9.13, which monitors the output. Perturbations  $\delta L_1$  and  $\delta L_2$  in the cavity lengths then produce a change

$$\delta\varphi_{o1} - \delta\varphi_{o2} \sim \frac{\omega}{\omega_{\text{gw}}} \frac{(\delta L_1 - \delta L_2)}{L} \quad (9.48)$$

in the relative phases at the beam splitter, and this in turn produces a change of the light power entering the photodetector. By using two cavities in this way, and keeping their light storage times (and hence response times) the same, one makes the light power entering the

photodetector be insensitive to fluctuations in the laser frequency; this is crucial for obtaining the high sensitivities that gravitational-wave detection requires.

The mirrors at the ends of each cavity are suspended as pendula, and when a gravitational wave with dimensionless amplitude  $h$  (to be discussed in Chap. 27) passes, it moves the mirrors back and forth, producing changes

$$\delta L_1 - \delta L_2 = hL \quad (9.49)$$

in the arm length difference. The resulting change in the relative phases of the two beams returning to the beam splitter,

$$\delta\varphi_{o1} - \delta\varphi_{o2} \sim \frac{\omega}{\omega_{\text{gw}}} h, \quad (9.50)$$

is monitored via the changes in power that it produces for the light going into the photodetector. If one builds the entire detector optimally and uses the best possible photodetector, these phase changes can be measured with a photon shot-noise-limited precision of  $\sim 1/\sqrt{N}$ . Here  $N \sim (W_\ell/\hbar\omega)(1/\omega_{\text{gw}})$  is the number of photons put into the detector by the laser (with power  $W_\ell$ ) during half a gravitational-wave period.<sup>8</sup> By combining this with Eq. (9.50) we see that the weakest wave that can be detected (at signal to noise ratio 1) is

$$h \sim \left( \frac{\hbar\omega_{\text{gw}}^3}{\omega W_\ell} \right)^{1/2}. \quad (9.51)$$

For a laser power  $W_\ell \sim 5$  Watts, and  $\omega_{\text{gw}} \sim 10^3\text{s}^{-1}$ ,  $\omega \sim 3 \times 10^{15}\text{s}^{-1}$ , this gravitational-wave sensitivity (noise level) is  $h \sim 3 \times 10^{-21}$ .

When operated in this manner, about 97 per cent of the light returns toward the laser from the beam splitter and the other 1 per cent goes out the end mirror or into the photodetector or gets absorbed or scattered due to imperfections in the optics. In LIGO's initial detectors, the 97 per cent returning toward the laser was recycled back into the interferometer, in phase with new laser light, by placing a mirror between the laser and the beam splitter. This “recycling mirror” (shown dashed in Fig. 9.13) made the entire optical system into a big optical resonator with two sub-resonators (the arms' Fabry-Perot cavities), and the practical result was a 50-fold increase in the input light power, from 5 Watts to 250 W—and an optical power in each arm of about  $\frac{1}{2} \times 250\text{W} \times 4/(1 - r_i^2) \sim 15$  kW; see Fig. 3 of LSC (2009). When operated in this manner, the interferometer achieved a sensitivity  $h \sim 3 \times 10^{-21}/\sqrt{50} \sim 4 \times 10^{-22}$ , which is close to the range expected for the strongest waves from colliding neutron stars, black holes, and other astrophysical sources; see Chap. 27. For a more accurate analysis of the sensitivity, see Exs. 9.18 and 9.19.

This estimate of sensitivity is actually the rms noise in a bandwidth equal to frequency at the minimum of LIGO's noise curve. Figure 6.7 in Chap. 6 shows the noise curve as the square root of the spectral density of the measured arm-length difference  $\xi \equiv L_1 - L_2$ ,  $\sqrt{S_\xi(f)}$ .

---

<sup>8</sup>This measurement accuracy is related to the Poisson distribution of the photons entering the interferometer's two arms: if  $N$  is the mean number of photons during a half gravitational-wave period, then the variance is  $\sqrt{N}$ , and the fractional fluctuation is  $1/\sqrt{N}$ . The interferometer's shot noise is actually caused by a beating of quantum electrodynamical vacuum fluctuations against the laser's light; for details see Caves (1980).

Since the waves produce a change of  $\xi$  given by  $\delta\xi = hL$  [Eq. (9.50)], the corresponding noise-induced fluctuations in the measured  $h$  have  $S_h = S_\xi/L^2$ , and the rms noise fluctuations in a bandwidth equal to frequency  $f$  are  $h_{rms} = \sqrt{S_h f} = (1/L)\sqrt{S_\xi f}$ . Inserting  $\sqrt{S_\xi} \simeq 10^{-19}$  m Hz $^{-1/2}$  and  $f \simeq 100$  Hz from Fig. 6.7, and  $L = 4$  km for the LIGO arm length, we obtain  $h_{rms} \simeq 3 \times 10^{-22}$ , in good agreement with our estimate above of  $4 \times 10^{-22}$ .

There are enormous obstacles to achieving such high sensitivity. To name just a few: Imperfections in the optics will absorb some of the high light power, heating the mirrors and beam splitter and causing them to deform. Even without such heating, the mirrors and beam splitter must be exceedingly smooth and near perfectly shaped to minimize the scattering of light from them (which causes noise; Ex. 8.17). Thermal noise in the mirrors and their suspensions (described by the fluctuation dissipation theorem) will cause the mirrors to move in manners that simulate the effects of a gravitational wave (Secs. 6.8.2 and 11.6.2), as will seismic- and acoustic-induced vibrations of the mirror suspensions. LIGO's arms must be long (4 km) in order to minimize the effects of these noises. While photon shot noise dominates near and above the noise curve's minimum,  $f \gtrsim 100$  Hz, these and other noises dominate at lower frequencies.

The initial LIGO detectors operated at their design sensitivity from autumn 2005 to autumn 2007, and then, after modest improvements, they operated again from spring 2009 to autumn 2010, carrying out gravitational-wave searches, much of the time in collaboration with international partners (the French-Italian VIRGO and British/German GEO600 interferometers). In autumn 2010, LIGO's detectors began a major, long-planned upgrade (to "advanced LIGO") that will increase their sensitivity ten-fold, bringing them into a range where they are likely to see a large number of gravitational wave sources, and carry out a rich program of observations. The result may be a revolution in our understanding of the universe.

\*\*\*\*\*

## EXERCISES

**Exercise 9.18** T2 *Derivation and Problem: Phase Shift in LIGO Arm Cavity*

In this exercise and the next, simplify the analysis by treating each Gaussian light beam as though it were a plane wave. The answers for the phase shifts will be the same as for a true Gaussian beam because, on the optic axis, the Gaussian beam's phase [Eq. (8.40a) with  $\varpi = 0$ ] is the same as that of a plane wave, except for the Gouy phase  $\tan^{-1}(z/z_0)$  which is very slowly changing and thus irrelevant.

- (a) For the interferometric gravitational wave detector depicted in Fig. 9.13 (with the arms' input mirrors having amplitude reflectivities  $\mathbf{r}_i$  close to unity and the end mirrors idealized as perfectly reflecting), analyze the light propagation in cavity 1 by the same techniques as were used for an etalon in Sec. 9.13. Show that, if  $\psi_{i1}$  is the light field impinging on the input mirror, then the total reflected light field  $\psi_{r1}$  is

$$\psi_{r1} = e^{i\varphi_1} \frac{1 - \mathbf{r}_i e^{-i\varphi_1}}{1 - \mathbf{r}_i e^{i\varphi_1}} \psi_{i1}, \quad \text{where} \quad \varphi_1 = 2kL_1. \quad (9.52a)$$

- (b) From this, infer that the reflected flux  $|\psi_{r1}|^2$  is identical to the cavity's input flux  $|\psi_{i1}|^2$ , as it must be since no light can emerge through the perfectly reflecting end mirror.
- (c) The arm cavity is operated on resonance, so  $\varphi_1$  is an integer multiple of  $2\pi$ . From Eq. (9.52a) infer that (up to fractional errors of order  $1 - \tau_i$ ) a change  $\delta L_1$  in the length of cavity 1 produces a change

$$\delta\varphi_{r1} = \frac{8k \delta L_1}{1 - \tau_i^2} . \quad (9.52b)$$

With slightly different notation, this is Eq. (9.47), which we used in the text's order of magnitude analysis of LIGO's sensitivity. In this exercise and the next, we will carry out a more precise analysis.

**Exercise 9.19** **T2** *Example: Photon Shot Noise in LIGO*

Continuing the preceding exercise, and continuing to treat the light beams as plane waves:

- (a) Denote by  $\psi_\ell$  the light field from the laser that impinges on the beam splitter and gets split in two, with half going into each arm. Using the above equations, infer that the light field returning to the beam splitter from arm 1 is  $\psi_{s1} = \frac{1}{\sqrt{2}}\psi_\ell e^{i\varphi_1}(1 + i\delta\varphi_{r1})$ , where  $\varphi_1$  is some net accumulated phase that depends on the separation between the beam splitter and the input mirror of arm 1.
- (b) Using the same formula for the field  $\psi_{s2}$  from arm 2, and assuming that the phase changes between beam splitter and input mirror are almost the same in the two arms, so  $\varphi_o \equiv \varphi_1 - \varphi_2$  is small compared to unity (mod  $2\pi$ ), show that the light field that emerges from the beam splitter, traveling toward the photodetector, is

$$\psi_{\text{pd}} = \frac{1}{\sqrt{2}}(\psi_{s1} - \psi_{s2}) = \frac{i}{2}(\varphi_o + \delta\varphi_{r1} - \delta\varphi_{r2})\psi_\ell \quad (9.53a)$$

to first order in the small phases. Show that the condition  $|\varphi_o| \ll 1$  corresponds to the experimenters' having adjusted the positions of the input mirrors in such a way that almost all of the light returns toward the laser and only a small fraction goes toward the photodetector.

- (c) For simplicity, let the gravitational wave travel through the interferometer from directly overhead and have an optimally oriented polarization. Then, as we shall see in Chap. 27, the dimensionless gravitational-wave field  $h(t)$  produces the arm-length changes  $\delta L_1 = -\delta L_2 = \frac{1}{2}h(t)L$ , where  $L$  is the unperturbed arm length. Show, then, that the field traveling toward the photodetector is

$$\psi_{\text{pd}} = \frac{i}{2}(\varphi_o + \delta\varphi_{gw})\psi_\ell , \quad \text{where} \quad \delta\varphi_{gw} = \frac{8kL}{1 - \tau_i^2}h(t) = \frac{16\pi L/\lambda}{1 - \tau_i^2}h(t) . \quad (9.53b)$$

The experimenter adjusts  $\varphi_o$  so it is large compared to the tiny  $\delta\varphi_{gw}$ .

- (d) Actually, this equation has been derived assuming, when analyzing the arm cavities [Eq. (9.52a)], that the arm lengths are static. Explain why it should still be nearly valid when the gravitational waves are moving the mirrors, so long as the gravitational-wave half period  $1/2f = \pi/\omega_{\text{gw}}$  is somewhat longer than the mean time that a photon is stored inside an arm cavity, i.e. so long as  $f \gg f_o$ , where

$$\boxed{f_o \equiv \frac{1 - \mathfrak{r}_i^2}{4\pi} \frac{c}{2L}}. \quad (9.54)$$

Assume that this is so. For the initial LIGO detectors,  $1 - \mathfrak{r}_i^2 \sim 0.03$  and  $L = 4$  km, so  $f_o \sim 90$  Hz.

- (e) Show that, if  $W_\ell$  is the laser power impinging on the beam splitter (proportional to  $|\psi_\ell|^2$ ), then the steady-state light power going toward the photodetector is  $W_{\text{pd}} = (\varphi_o/2)^2 W_\ell$  and the time-variation in that light power due to the gravitational wave (the gravitational-wave signal) is

$$W_{\text{gw}}(t) = \sqrt{W_\ell W_{\text{pd}}} \frac{16\pi L/\lambda}{1 - \mathfrak{r}_i^2} h(t). \quad (9.55a)$$

The photodetector monitors these changes  $W_{\text{gw}}(t)$  in the light power  $W_{\text{pd}}$  and from them infers the gravitational-wave field  $h(t)$ . This is called a “DC” or “homodyne” readout system; it works by beating the gravitational-wave signal field ( $\propto \delta\varphi_{\text{GW}}$ ) against the steady light field (“local oscillator”,  $\propto \varphi_o$ ) to produce the signal light power  $W_{\text{gw}}(t) \propto h(t)$ .

- (f) Shot noise in the interferometer’s output light power  $W_{\text{pd}}$  gives rise to noise in the measured gravitational-wave field  $h(t)$ . From Eq. (9.55a) show that the spectral density of the noise in the measured  $h(t)$  is

$$S_h(f) = \left( \frac{(1 - \mathfrak{r}_i^2)\lambda}{16\pi L} \right)^2 \frac{S_{W_{\text{pd}}}}{W_\ell W_{\text{pd}}}. \quad (9.55b)$$

In Sec. 6.7.4, we derived the formula  $S_{W_{\text{pd}}} = 2\mathcal{R}(\hbar\omega)^2 = 2W_{\text{pd}}\hbar\omega$  for the (frequency-independent) spectral density of a steady, monochromatic light beam’s power fluctuations due to shot noise; here  $\mathcal{R} = W_{\text{pd}}/\hbar\omega$  is the average rate of arrival of photons. Combining with Eq. (9.55b), deduce your final formula for the spectral density of the noise in the inferred gravitational-wave signal

$$S_h(f) = \left( \frac{(1 - \mathfrak{r}_i^2)\lambda}{16\pi L} \right)^2 \frac{2}{W_\ell/\hbar\omega}; \quad (9.56a)$$

and from this deduce the rms noise in a bandwidth equal to frequency

$$\boxed{h_{\text{rms}} = \sqrt{f S_h} = \left( \frac{(1 - \mathfrak{r}_i^2)\lambda}{16\pi L \sqrt{N}} \right)}, \quad \text{where} \quad \boxed{N = \frac{W_\ell}{\hbar\omega} \frac{1}{2f}} \quad (9.56b)$$

is the number of photons that impinge on the beam splitter, from the laser, in half a gravitational-wave period.

- (g) In the next exercise we shall derive (as a challenge) the modification to the spectral density that arises at frequencies  $f \gtrsim f_o$ . The signal strength that gets through the interferometer is reduced because the arm length is increasing, then decreasing, then increasing again, ... while the typical photon is in an arm cavity. The result of the analysis is an increase of  $S_h(f)$  by  $1 + (f/f_o)^2$ , so

$$S_h(f) = \left( \frac{(1 - \tau_i^2)\lambda}{16\pi L} \right)^2 \frac{2}{W_\ell/\hbar\omega} \left( 1 + \frac{f^2}{f_o^2} \right). \quad (9.57)$$

Compare this with the measured noise, at frequencies above  $f_o \sim 90$  Hz in the initial-LIGO detectors (Fig. 6.7 with  $\xi = hL$ ), using the initial-LIGO parameters,  $\lambda = 1.06\mu\text{m}$ ,  $\omega = 2\pi c/\lambda \simeq 2 \times 10^{15} \text{ s}^{-1}$ ,  $L = 4 \text{ km}$ ,  $W_\ell = 150 \text{ W}$ ,  $1 - \tau_i^2 = 1/30$ . It should agree fairly well with the measured noise at frequencies  $f \gtrsim f_o$  where most of the noise is due to photon shot noise. Also compare the noise (9.57) in a bandwidth equal to frequency,  $\sqrt{f S_h}$ , evaluated at frequency  $f = f_o$ , with the crude estimate (9.51) worked out in the text. They should agree to within a factor of order unity.

**Exercise 9.20** T2 *Challenge: LIGO Shot Noise at  $f \gtrsim f_o$*

Derive the factor  $1 + (f/f_o)^2$  by which the spectral density of the shot noise is increased at frequencies  $f \gtrsim f_o$ . [Hint: Redo the analysis of the arm cavity fields, part (a) of Ex. 9.19 using an arm length that varies sinusoidally at frequency  $f$  due to a sinusoidal gravitational wave, and then use the techniques of Ex. 9.19 to deduce  $S_h(f)$ .]

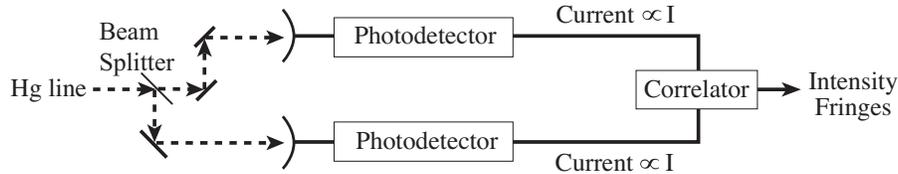
\*\*\*\*\*

## 9.6 Power Correlations and Photon Statistics: Hanbury Brown & Twiss Intensity Interferometer

A type of interferometer that is rather different from those studied above was proposed and constructed by Robert Hanbury Brown and Richard Q. Twiss in 1956. In this interferometer, light powers rather than amplitudes are combined to measure the degree of coherence of the radiation field. This is often called an *intensity interferometer* because the optics community often uses the word “intensity” to mean energy flux (power per unit area).

In their original experiment, Hanbury Brown and Twiss divided light from an incandescent mercury lamp and sent it along two paths of variable length before detecting photons in each beam separately using a photodetector; see Fig. 9.14. The electrical output from each photodetector measures the rate of arrival of its beam’s photons, or equivalently its beam’s power  $W(t)$ , which we can write as  $K(\Re\Psi)^2$ , where  $K$  is a constant. This  $W$  exhibits fluctuations  $\delta W$  about its mean value  $\overline{W}$ , and it was found that the fluctuations in the two beams were correlated. How can this be?

The light that was detected originated from many random and independent emitters and therefore obeys Gaussian statistics, according to the central limit theorem (Sec. 6.3.2). This



**Fig. 9.14:** Hanbury Brown and Twiss intensity interferometer.

turns out to mean that the fourth-order correlations of the wave field  $\Psi$  with itself can be expressed in terms of the second-order correlations—i.e., in terms of the degree of coherence  $\gamma_{||}$ . More specifically:

Continuing to treat the wave field  $\Psi$  as a scalar, (i) we write each beam's power as the sum over a set of Fourier components  $\Psi_j$  with precise frequencies  $\omega_j$  and slowly wandering, complex amplitudes  $W(t) = (\sum_j \Re \Psi_j)^2$ , (ii) we form the product  $W(t)W(t + \tau)$ , (iii) we keep only terms that will have nonzero averages by virtue of containing products of the form  $\propto e^{+i\omega_j t} e^{-i\omega_j t} e^{+i\omega_k t} e^{-i\omega_k t}$  (where  $j$  and  $k$  are generally not the same), and we average over time. Thereby we obtain

$$\begin{aligned} \overline{W(t)W(t + \tau)} &= K^2 \overline{\Psi(t)\Psi^*(t)} \times \overline{\Psi(t + \tau)\Psi^*(t + \tau)} + K^2 \overline{\Psi(t)\Psi^*(t + \tau)} \times \overline{\Psi^*(t)\Psi(t + \tau)} \\ &= \overline{W}^2 [1 + |\gamma_{||}(\tau)|^2] \end{aligned} \quad (9.58)$$

[cf. Eq. (9.16) with  $\Psi$  allowed to be complex]. If we now measure the relative fluctuations, we find that

$$\frac{\overline{\delta W(t)\delta W(t + \tau)}}{\overline{W(t)}^2} = \frac{\overline{W(t)W(t + \tau)} - \overline{W(t)}^2}{\overline{W(t)}^2} = |\gamma_{||}(\tau)|^2. \quad (9.59)$$

[Note: This analysis is only correct if the radiation comes from many uncorrelated sources—the many independently emitting mercury atoms in Fig. 9.14—and therefore has Gaussian statistics.]

Equation (9.59) tells us that the power as well as the amplitude of coherent radiation must exhibit a positive longitudinal correlation; and the degree of coherence for the fluxes is equal to the squared modulus of the degree of coherence for the amplitudes. Although this result was rather controversial at the time the experiments were first performed, it is easy to interpret qualitatively if we think in terms of photons rather than classical waves. Photons are bosons and are therefore positively correlated even in thermal equilibrium; cf. Chaps. 3 and 4. When they arrive at the beam splitter of Fig. 9.14, they clump more than would be expected for a random distribution of classical particles. In fact, treating the problem from the point of view of photon statistics gives an answer equivalent to Eq. (9.59).

Some practical considerations should be mentioned. The first is that our result (9.59), derived for a scalar wave, is really only valid for electromagnetic waves if they are completely polarized. If the incident waves are unpolarized, then the intensity fluctuations are reduced by a factor two. The second point is that, in the Hanbury Brown and Twiss experiments, the photon counts were actually averaged over longer times than the correlation time of the incident radiation. This reduced the magnitude of the measured effect further.

Nevertheless, after successfully measuring temporal power correlations, Hanbury Brown and Twiss constructed a stellar intensity interferometer with which they were able to measure

the angular diameters of bright stars. This method had the advantage that it did not depend on the phase of the incident radiation, so the results were insensitive to atmospheric fluctuations (seeing), one of the drawbacks of the Michelson stellar interferometer (Sec. 9.2.5). Indeed, it is not even necessary to use accurately ground mirrors to measure the effect. The method has the disadvantage that it can only measure the modulus of the degree of coherence; the phase is lost. It was the first example of using fourth-order correlations of the light field to extract image information from light that has passed through the earth's turbulent atmosphere (Box 9.2).

\*\*\*\*\*

## EXERCISES

### Exercise 9.21 *Derivation: Power Correlations*

By expressing the field as either a Fourier sum or a Fourier integral complete the argument outlined in Eq. (9.58).

### Exercise 9.22 *Problem: Electron Intensity Interferometry*

Is it possible to construct an intensity interferometer (i.e., a number flux interferometer) to measure the coherence properties of a beam of electrons? What qualitative differences do you expect there to be from a photon intensity interferometer? What do you expect Eq. (9.59) to become?

\*\*\*\*\*

## Bibliographic Note

For pedagogical introductions to interference and coherence at an elementary level, with greater detail than this chapter, see Hecht (2002) and Klein & Furtak (1986). For more advanced treatments, we like Ghatak (2010), Saleh and Teich (2007), Pedrotti (2007), and especially Brooker (2003). For an particularly deep and thorough discussion of coherence, see Goodman (1985). For modern applications of interferometry (including the optical frequency comb), see Yariv and Pochi (2007), and at a more elementary level, Hariharan (2007).

## Bibliography

Black, Eric D. 2001. "An introduction to Pound-Drever-Hall laser frequency stabilization", *American Journal of Physics* **69**, 79–87.

Brooker, Geoffrey 2003. *Modern Classical Optics*, Oxford: Oxford University Press

Caves, C. M. 1980 "Quantum-mechanical radiation pressure fluctuations in an interferometer", *Physical Review Letters*, **45**, 75.

- Cundiff, Steven T. 2002. “Phase stabilization of ultrashort optical pulses”, *Journal of Physics D: Applied Physics* **35**, r43–R59.
- Cundiff, Steven T. and Ye, Jun 2003. “Colloquium: Femtosecond optical frequency combs”, *Reviews of Modern Physics* **75**, 325–342.
- Feynman, R. P., Leighton, R. B., & Sands, M. 1965 *The Feynman Lectures on Physics*, New York: Addison Wesley
- Francon, M. 1966 *Optical Interferometry* New York: Academic Press
- Ghatak, Ajoy 2010. *Optics*, New York: McGraw-Hill
- Goodman, Joseph W. 1985. *Statistical Optics*, New York: Wiley
- Hariharan, P. 2007. *Basics of Interferometry*, second edition, New York: Academic Press
- Hecht, E. 2002 *Optics*, fourth edition, New York: Addison Wesley
- Klein, Miles V. & Furtak, Thomas E. 1986 *Optics* New York: Wiley
- LSC: The LIGO Scientific Collaboration, 2009. “LIGO: The Laser Interferometer Gravitational-Wave Observatory”, *Reports on Progress in Physics* **72**, 076901.
- Mather, J. C. et al 1994 “Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument”, *Astrophysical Journal* **420**, 439.
- Pedrotti, Frank L., Pedrotti, Leno S., and Pedrotti, Leno M. 2007. *Introduction to Optics*, third edition, Upper Saddle River, NJ: Pearson
- Rodier, F. 1981. “The effects of atmospheric turbulence in optical astronomy”, *Progress in Optics*, **19**, 281–376.
- Saleh, B. E. A. and Teich, M. C. 2007. *Fundamentals of Photonics*, New York: Wiley
- Yariv, Amnon and Yeh, Pochi 2007. *Photonics: Optical Electronics in Modern Communications*, Sixth Edition, Oxford: Oxford University Press.

### Box 9.4 Important Concepts in Chapter 9

- Interference Fringes – Sec. 9.2.1 and Fig. 9.1
- Incoherent radiation – Eqs. (9.4) and (9.5)
- Degrees of Coherence and Fringe Visibility
  - Degree of lateral coherence (complex fringe visibility) for nearly monochromatic radiation,  $\gamma_{\perp}$  – Eqs. (9.6a), (9.10) and (9.12); and discussion after Eq. (9.8)
  - Visibility for lateral coherence:  $V = |\gamma_{\perp}|$  – Eq. (9.8)
  - Degree of temporal (or longitudinal) coherence for nearly monochromatic radiation – Eq. (9.15)
  - Degree of temporal coherence for broad-band radiation – Eq. (9.16)
  - Three-dimensional degree of coherence – Sec. 9.2.8
- Coherence lengths and times – Eqs. (9.13), (9.17) and associated discussions, and passage following Eq. (9.21)
- van Cittert-Zernike Theorem relating degree of coherence to angular distribution and/or spectrum of the source
  - For lateral coherence, Eqs. (9.8) and (9.17)
  - For temporal coherence of broad-band radiation – Eqs. (9.18)
  - Three dimensional (lateral and longitudinal together) – Eqs. (9.22)
  - Relationship to Wiener-Khintchine theorem – Ex. (9.7b)
- Michelson stellar interferometer and **T2** astronomical seeing, – Sec. 9.2.5 and Box 9.2
- Michelson interferometer and Fourier-transform spectroscopy — Fig. 9.3, Sec. 9.2.7
- Complex random processes – Ex. 9.8
- Radio Telescope: How one constructs images of the source, and what determines its angular resolution – Sec. 9.3
- Amplitude reflection and transmission coefficients – Eq. (9.30a)
- Reciprocity relations for reflection and transmission coefficients– Eqs. (9.32), Ex. 9.10
- High reflectivity coatings and anti-reflection coatings constructed from alternating dielectric layers – Exs. 9.10 (first paragraph) and 9.13
- Etalon and Fabry-Perot interferometer – Secs. 9.4
  - Finesse and its influence on half-width of resonance, phase shift across resonance, and response time – Eqs. (9.38)–(9.40) and Fig. 9.9.
  - Free spectral range – passage following Eq. (9.38)
  - Some applications of Fabry-Perot interferometers, especially PDH laser locking and the optical frequency comb – Sec. 9.4.2
- Sagnac interferometer – Ex. 9.17
- **T2** Laser interferometer gravitational-wave detector, and how it works – Sec. 9.5
- Light-power correlations and intensity interferometry, Sec. 9.6